

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2003, 14:38:21 : Search time 2327 seconds

(without alignments)  
12527.665 Million cell updates/sec

Title: US-09-319-724A-13

Perfect score: 1800

Sequence: 1 ccgcgcgcgtgtgctgtg.....gtggaacttgcacagcaaaa 1800

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: gb\_gss:\*

18: em\_gss\_hum:\*

19: em\_gss\_inv:\*

20: em\_gss\_pln:\*

21: em\_gss\_vrt:\*

22: em\_gss\_fun:\*

23: em\_gss\_mam:\*

24: em\_gss\_mus:\*

25: em\_gss\_Other:\*

26: em\_gss\_pro:\*

27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	812.6	45.1	879	14	BQ437992
2	793.2	44.1	1098	13	BM554832
3	734.4	40.8	1121	13	BM466309
4	707.6	39.3	1046	13	BM462090
5	693.6	38.5	744	12	BF732416
6	683.8	38.0	976	14	BM804893

7	663.6	36.9	700	12	BF110352
8	663.2	36.8	682	12	BF445098
9	655.4	36.4	749	9	AI317419
10	653.4	36.3	663	12	BF062374
11	652.2	36.2	699	12	BG252894
12	648.8	36.0	693	12	BG499066
13	642.8	35.7	748	13	BI089285
14	634.8	35.3	752	10	AW475143
15	618.4	34.4	772	12	BG252656
16	614.4	34.1	619	10	AW576585
17	613.8	34.1	714	10	AW475466
18	610	33.9	918	12	BG287562
19	601.2	33.4	670	12	BG621630
20	598	33.2	939	14	BQ958528
21	590.2	32.8	971	12	BG171028
22	583.4	32.4	608	12	BF941584
23	583	32.4	616	14	BQ638442
24	581.2	32.3	721	12	BE868698
25	578	32.1	684	12	BE848488
26	577.6	32.1	601	10	AW305358
27	576.4	32.0	912	12	BE905485
28	575.4	32.0	662	9	AI172333
29	567.8	31.5	731	13	BI868308
30	564.4	31.4	566	14	BM827392
31	563.4	31.3	610	10	AW263003
32	559.4	31.1	574	12	BG492937
33	559.4	31.1	693	9	AA057428
34	558	31.0	587	12	BF434917
35	551.6	30.6	848	12	BG244043
36	548.8	30.5	681	12	BF105260
37	546.4	30.4	709	9	AA800019
38	545.2	30.3	769	12	BG619729
39	542	30.1	547	12	BG718166
40	538.2	29.9	903	14	BQ233844
41	536	29.8	552	9	AL550770
42	535.4	29.7	926	12	BE872774
43	534.6	29.7	559	12	BF447674
44	534.6	29.7	852	12	BF181943
45	531.6	29.5	960	13	BI852732

ALIGNMENTS

RESULT 1  
LOCUS BQ437992 879 bp mRNA linear EST 24-MAY-2002  
DEFINITION AGENCOURT\_7894720 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6158457  
5', mRNA sequence.  
ACCESSION BQ437992  
VERSION BQ437992.1 GI:21177068  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 879)  
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLNMI3505 row: p column: 10  
High quality sequence stop: 627.  
Location/Qualifiers 1. .879  
source

[illegible]

Db	300	TATTTAGACCATCAAGTCACCCACTGATATTTCTCTCTGATGGTTCTGGATGTCAGA	359
Qy	865	TATTTGCTGTCTCTCATCGTTATTTATTTGCAATGATAGAAGATTTATATACTGAGA	924
Db	360	TATTTGCTGTCTCTCATCGTTATTTATTTGCAATGATAGAAGATTTATATACTGAGA	419
Qy	925	GGGGATCAATGCTCAGTAGACAGCCATATTTGTCATGCTACGTCTCCAGTGAATGGTT	984
Db	420	GGGGATCAATGCTCAGTAGACAGCCATATTTGTCATGCTACGTCTCCAGTGAATGGTT	479
Qy	985	ATTTTGGAGGAAGTCTGTATGCTACAGAAAGGAGAGATGGATAAAGACAGATGTTTA	1044
Db	480	ATTTTGGAGGAAGTCTGTATGCTACAGAAAGGAGAGATGGATAAAGACAGATGTTTA	539
Qy	1045	TTGGGGCATTCCTTATCCCAAGTATGGTGTGTGGACATTCCTTTCATCAATTTTCATAG	1104
Db	540	TTGGGGCATTCCTTATCCCAAGTATGGTGTGTGGACATTCCTTTCATCAATTTTCATAG	599
Qy	1105	CCATTTATTAACCATGCTTCAAGAGCCATTCCCTTTTGGAAACAATGGTGCCCGTTGTTGCA	1164
Db	600	CCATTTATTAACCATGCTTCAAGAGCCATTCCCTTTTGGAAACAATGGTGCCCGTTGTTGCA	659
Qy	1165	TCCTGTTTTTTGTTATTCCTTCCTCTAAATCTGTGTGGTACAACTACTGGCCGAAATCTGT	1224
Db	660	TCCTGTTTTTTGTTATTCCTTCCTCTAAATCTGTGTGGTACAACTACTGGCCGAAATCTGT	719
Qy	1225	CAGTGTAGCCCAAC-TTTTCCTTGTGCTGTCAATGCTGTGCTCGTCCCTATACC-GGAGAA	1282
Db	720	CAGTGTAGCCCAAC-TTTTCCTTGTGCTGTCAATGCTGTGCTCGTCCCTATACC	779
Qy	1283	AAATATGGTTTCATGGAGCTCGGGTTATTTGTTTGCTGGG--TGGAAATTTTACCTTTGGT	1340
Db	780	AAATATGGTTTCATGGAGCTCGGGTTATTTGTTTGCTGGGTTGGAAATTTTACCTTTGGT	839
Qy	1341	TCAATCTTTTATTTGAAA	1356
Db	840	TCAATCTTTTATTTGAA	855

RESULT 3					
BM466309					
LOCUS	BM466309	1121 bp	mRNA	linear	EST 05-FEB-2002
DEFINITION	AGENCOURT_6457073 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:5577041 5', mRNA sequence.				

**SOURCE:** Human.  
**ORGANISM:** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1121)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC).  
Unpublished (1999)  
**REFERENCE:**  
Contact: Robert Strausberg, Ph.D.  
Email: [cqabbs-r@mail.nih.gov](mailto:cqabbs-r@mail.nih.gov)

## FEATURES

```

/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
341 a 211 c 250 q 319 t
BASE COUNT

```

Query Match	40.8%	Score 734.4	DB 13	Length 1121
Best Local Similarity	97.6%	Pred. No. 2e-169		
Matches 777	Conservative 0	Mismatches 16	Indels 3	Gaps 3
Qy 60	CACAGCTATCAAGATAAAGAGGAAGTTCCTTATGCGATCAATACACTGTTGGGCCCTTACCAT	119		
Db 74	CAGCAGTATCAAGATAAAGAGGAAGTTCCTTATGCGATCAATACACTGTTGGGCCCTTACCAT	133		
Qy 120	AATCGTCAAGAAACATATAAGTACTTTTTCACCTTCCATTCGTGTGGGTCAAAAAAAGT	179		
Db 134	AATCGTCAAGAAACATATAAGTACTTTTTCACCTTCCATTCGTGTGGGTCAAAAAAAGT	193		
Qy 180	ATCAGTCATTAACCATGAAACTCTGGGAGAAGCATTCAAGGGTTGAAATGGAAATTTAGT	239		
Db 194	ATCAGTCATTAACCATGAAACTCTGGGAGAAGCATTCAAGGGTTGAAATGGAAATTTAGT	253		
Qy 240	GGTCTGGATATTAATTTTAAAGATGATGTGATGCCAGCCACTTACTGTGAAATTTGATTTA	299		
Db 254	GGTCTGGATATTAATTTTAAAGATGATGTGATGCCAGCCACTTACTGTGAAATTTGATTTA	313		
Qy 300	GATAAGAAAAAGAGATGCATTTGTATATGCCATAAAAAATCATTTACTGGTACCAGATG	359		
Db 314	GATAAGAAAAAGAGATGCATTTGTATATGCCATAAAAAATCATTTACTGGTACCAGATG	373		
Qy 360	TACATAGATGATTTACCAATATGGGGTATTGTGGTGAGGTGATGAAAAATGGAGAAGAT	419		
Db 374	TACATAGATGATTTACCAATATGGGGTATTGTGGTGAGGTGATGAAAAATGGAGAAGAT	433		
Qy 420	TACTATCTTTTGGACCTATATAAAAACTTCAAAATAGTTTTAAATGGAAATCGAAATGCTTGCAT	479		
Db 434	TACTATCTTTTGGACCTATATAAAAACTTCAAAATAGTTTTAAATGGAAATCGAAATGCTTGCAT	493		
Qy 480	GTTAATCTTAAGTCTAGTGAAGGAAAGTGAAACTGGTTCCAAATCTATAAATCCAGATGTCA	539		
Db 494	GTTAATCTTAAGTCTAGTGAAGGAAAGTGAAACTGGTTCCAAATCTATAAATCCAGATGTCA	553		
Qy 540	TATTTCAGTAAATGGAAAAAGTCAGATGTAATTTGAAGATCGAATTTGACAAATATCTT	599		
Db 554	TATTTCAGTAAATGGAAAAAGTCAGATGTAATTTGAAGATCGAATTTGACAAATATCTT	613		
Qy 600	GATCCGTCTTTTTCACATCGGATTCATTTGCTTTTCAATTTTCAACTCTCTTCATGATG	659		
Db 614	GATCCGTCTTTTTCACATCGGATTCATTTGCTTTTCAATTTTCAACTCTCTTCATGATG	673		
Qy 660	GTGATCTTCTTGGTGGGCTTATGTTTCAATGATTTTAAATGAGAACATTAAGAAAAAGATTAT	719		
Db 674	GTGATCTTCTTGGTGGGCTTATGTTTCAATGATTTTAAATGAGAACATTAAGAAAAAGATTAT	733		
Qy 720	GCTCGGTACAGTAAGAGGAAAGAAATGGATGATATGGATAGAGACCTAGG-AGATGAATA	778		
Db 734	GCTCGGTACAGTAAGAGGAAAGAAATGGATGATATGGATAGAGACCTAGGAAACAGGAATA	793		
Qy 779	TGGATGGAACAGGTGAT-CCGATGTATTTAGACCATCAAGTCACCCACTG-ATATTTT	836		
Db 794	TGAATGGAACAGGGGCTGGGAGATGTATTTAGACCATCAAGTCCCCCTGGAATATTT	853		
Qy 837	TCCCTCTCTGATTTGGTT	852		
Db 854	TCCCTCCCGGATGGT	869		

RESULT 4

BM462090

LOCUS	BM462090	1046 bp	mRNA	linear	EST 05-FEB-2007
DEFINITION	AGENCOURT_6424576 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5491167				
	5', mRNA sequence.				
ACCESSION	BM462090				
VERSION	BM462090.1	GI:18511130			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .				
TITLE	1 (bases 1 to 1046)				
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)				
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapsb-femail.nih.gov">cgapsb-femail.nih.gov</a> Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LLAM12111 row: d column: 16 High quality sequence stop: 574.				
FEATURES	Location/Qualifiers				
source	1..1046 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5491167" /clone_lib="NIH_MGC_67" /tissue_type="retinoblastoma" /lab_host="DH10B (phage-resistant)" /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.75 Kb. Library constructed by Life Technologies."				
BASE COUNT	272 a	209 c	205 g	360 t	
ORIGIN					
Query Match	39.38;	Score	707.6;	DB 13;	Length 1046;
Best Local Similarity	99.38;	Pred.	No. 7.6e-163;		
Matches	721;	Conservative	0;	Mismatches	4;
				Indels	1;
				Gaps	1;
QY	1076	TGGCACTGCCTCTTCATCAATTTTCATAGCCATTTATTACCATGCTTCAAGAGCCATTC	1135		
Db	1	TGGCACTGCCTCTTCATCAATTTTCATAGCCATTTATTACCATGCTTCAAGAGCCATTC	60		
QY	1136	TTTTTGGACAATGTCGCCGTTGTCATCTGTTTTTTCGTTTATCTCTCTCTAAATCT	1195		
Db	61	TTTTTGGACAATGTCGCCGATTGTGCATCTGTTTTTTCGTTTATCTCTCTCTAAATCT	120		
QY	1196	TGTTGGTACAATACCTTGCCCGAAATCTGTCAAGTCAGCCCAACTTTCCTTCTCGTGTCAA	1255		
Db	121	TGTTGGTACAATACCTTGCCCGAAATCTGTCAAGTCAGCCCAACTTTCCTTCTCGTGTCAA	180		
QY	1256	TGCTGTCCTGTCCTATACCGGAGAAAAATGGTTTCATGGAGCGTCGGGTTATTGTTG	1315		
Db	181	TGCTGTCCTGTCCTATACCGGAGAAAAATGGCTCATGGAGCGTCGGGTTATTGTTG	240		
QY	1316	CCTGGGTGGAAATTTTACCTTTTGGTTCAACTCTTATTTAGAAATGATTTTCATCTTCACGTC	1375		
Db	241	CCTGGGTGGAAATTTTACCTTTTGGTTCAACTCTTATTTGAAATGATTTTCATCTTCACGTC	300		
QY	1376	TTTCTGGGCATATAAGATCTATTATGTCTATGGCTTCATGATGCTGGTGTGTTATCCT	1435		
Db	301	TTTCTGGGCATATAAGATCTATTATGTCTATGGCTTCATGATGCTGGTGTGTTATCCT	360		
QY	1436	GTGCATTGTGACTGTCGTGTGACTATTGTGCACATATTTTCTACTAAATGAGAGA	1495		
Db	361	GTGCATTGTGACTGTCGTGTGACTATTGTGTGCACATATTTTCTACTAAATGAGAGA	420		
QY	1496	TTACCGGTGGCAATGGACAAAGTTTTTCTCTCTGTCGTCATCAACTGCAATCTATGTTTACAT	1555		

Db	421	TTACCGGTGGCAATGGACAAGTTTTCTCTCUGCTGCATCAACTGCAATCTATGTTTACAT	480
Qy	1556	GTAFTCCCTTTTACTACATATTTTTTCACAAACAAAAGATGTATGGCTTATTTCCAAACATCATTT	1615
Db	481	GTAFTCCCTTTTACTACATATTTTTTCACAAACAAAAGATGTATGGCTTATTTCCAAACATCATTT	540
Qy	1616	TTACTTTTGGATATATGCGGCTATTTAGCACAGCCTTGCGGATAATGTGTGGAGCGAATGG	1675
Db	541	TTACTTTTGGATATATGCGGCTATTTACCACAGCCTTGCGGATAATGTGTGGAGCGAATGG	600
Qy	1676	TTACATGGGAACAAGTGCCTTTTGTGCCAAAAAATCTATATAATGTGAAAAATTGACTAGAG	1735
Db	601	TTACATGGGAACAAGTGCCTTTTGTGCCAAAAAATCTATATAATGTGAAAAATTGACTAGAG	660
Qy	1736	ACCCAAGAAAACCTGGGAACCTTTTGATCAATTTCTTTTTCATAGGGGT-GGAACCTTGACA	1794
Db	661	ACCCAAGAAAACCTGGGAACCTTTTGATCAATTTCTTTTTCATAGGGGTGGGAACCTTGACA	720
Qy	1795	GCAAAA 1800	
Db	721	GCAAAA 726	
RESULT 5	BF732416	BF732416	744 bp mRNA linear EST 30-MAR-2001
LOCUS	nae09f02.x1	NCI_CGAP_Ov18 Homo sapiens cDNA clone IMAGE:3435002 3,	
DEFINITION	similar to TR:Q9UHW8 Q9UHW8 SM-I1044 BINDING PROTEIN ;, mRNA sequence.		
ACCESSION	BF732416	GI:12057419	
VERSION	EST.		
KEYWORDS	Homo sapiens		
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 744)		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index		
JOURNAL	Unpublished (1997)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: ccapbs-remail.nhl.gov Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lnl.gov Seq primer: -400P from Gibco High quality sequence stop: 495.		
FEATURES	Location/Qualifiers		
source	1..744		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3435002"		
	/clone_lib="NCI_CGAP_Ov18"		
	/tissue_type="fibrotheoma"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: ovary; Vector: pT7M3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'; TGTTACCAATCTGAATGGAGCGCGCGACATTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7M3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."		
BASE COUNT	243 a 120 c 161 g 220 t		
ORIGIN			











```
Db 661 GGC 663
||||
RESULT 11
BC252894
LOCUS
DEFINITION
  BG252894 602365488F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4473861 5',
  mRNA sequence.
ACCESSION
  BG252894
VERSION
  BG252894.1 GI:12762710
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 699)
  NIH-MGC http://mgc.nci.nih.gov/.
  National Institutes of Health, Mammalian Gene Collection (MGC)
  Unpublished (1999)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  Tissue Procurement: ATCC
  cDNA Library Preparation: Life Technologies, Inc.
  CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be
  found through the I.M.A.G.E. Consortium/LLNL at:
  http://image.llnl.gov
  Plate: LLAM10296 row: p column: 22
  High quality sequence stop: 697.
  Location/Qualifiers
    1..699
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone_lib="NIH_MGC_90"
    /tissue_type="adenocarcinoma, cell line"
    /lab_host="DH10B (phage-resistant)"
    /note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
    Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
    Average insert size 1.7 Kb. Library enriched for
    full-length clones and constructed by Life Technologies.
    Note: this is a NIH_MGC Library."
BASE COUNT 168 a 133 c 140 g 258 t
ORIGIN

Query Match 36.2%; Score 652.2; DB 12; Length 699;
Best Local Similarity 97.7%; Pred. No. 2.9e-149;
Matches 683; Conservative 0; Mismatches 13; Indels 3; Gaps 2;

QY 1084 CCTTCTTCATCAATTCATAGCCATTATTACCATGCTTCAAGAGCCATTCCTTTGGAA 1143
Db 1 CCTTCTTCATCAATTCATAGCCATTATTACCATGCTTCAAGAGCCATTCCTTTGGAA 60

QY 1144 CAATGGTGGCCGTTTGTTCATCTGTTTTTTTGTATTCTCTCTAAATCTTTGTGGTA 1203
Db 61 CAATGGTGGCCGTTTGTTCATCTGTTTTTTTGTATTCTCTCTAAATCTTTGTGGTA 120

QY 1204 CAATACTTGGCCGAAATCTGTCTAGGTGAGCCCAACTTTCCTTGTCTGTGTCATGCTGTGC 1263
Db 121 CAATACTTGGCCGAAATCTGTCTAGGTGAGCCCAACTTTCCTTGTCTGTGTCATGCTGTGC 180

QY 1264 CTGCTCCTACCGGAGAAAAATGTTTCATGAGCGCTCGGGTATTGTTGCTGGGTG 1323
Db 181 CTGCTCCTACCGGAGAAAAATGTTTCATGAGCGCTCGGGTATTGTTGCTGGGTG 240

QY 1324 GAATTTTACCTTTTGGTTTCAATCTTTTATTGAAATGATTTCATCTTCACCTCTTTCTGGG 1383
Db 241 GAATTTTACCTTTTGGTTTCAATCTTTTATTGAAATGATTTCATCTTCACGCTCTTTCTGGG 300

QY 1384 CATATAAGATCATATTATGCTATFGGCTTCATGATCGTGGTGGTATTCCTGTGCATTTG 1443
|||||
```





DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10298 row: j column: 03  
High quality sequence stop: 702.

FEATURES

Location/Qualifiers  
1..772  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4474466"  
/clone\_lib="NIH\_MGC\_90"  
/tissue\_type="adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: liver; Vector: pCMV-SPORT6; Site:1: NotI;  
Site:2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 Kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."  
BASE COUNT 186 a 144 c 164 g 278 t  
ORIGIN

Query Match 34.4%; Score 618.4; DB 12; Length 772;  
Best Local Similarity 95.8%; Pred. No. 5.8e-141;  
Matches 711; Conservative 0; Mismatches 21; Indels 10; Gaps 7;

QY 1049 GGCATTCCCTATCCAGCTATGGTGTGGCACTGCCTTCTTCATCAATTCATAGCCAT 1108  
Db 1 GGCATTCCCTATCCAGCTATGGTGTGGCACTGCCTTCTTCATCAATTCATAGCCAT 60  
QY 1109 TTATTACCATGCTTCAAGAGCCATTCCCTTTTGGACAATGGTGGCCGTTTGTGCATCTG 1168  
Db 61 TTATTACCATGCTTCAAGAGCCATTCCCTTTTGGACAATGGTGGCCGTTTGTGCATCTG 120  
QY 1169 TTTTGTGTTATTCCTCCTAAATCTTGTGTGTAATACCTTGGCCGAAATCTGTCTAGG 1228  
Db 121 TTTTGTGTTATTCCTCCTAAATCTTGTGTGTAATACCTTGGCCGAAATCTGTCTAGG 180  
QY 1229 TCAGCCCAACTTTCCCTGCTGCTCAATGCTGCTGCTGCTATACCGGAGAAAAAATG 1288  
Db 181 TCAGCCCAACTTTCCCTGCTGCTCAATGCTGCTGCTGCTATACCGGAGAAAAAATG 240  
QY 1289 GTTCATGGAGCCTCGCGGTTATTGTTGCTGGTGGAAATTTACCTTTTGGTTCAATCTT 1348  
Db 241 GTTCATGGAGCCTCGCGGTTATTGTTGCTGGTGGAAATTTACCTTTTGGTTCAATCTT 300  
QY 1349 TATTGAATGTAATTCATCTTCACGCTCTTCGCGCATATAGATCTATATGCTCTATGG 1408  
Db 301 TATTGAATGTAATTCATCTTCACGCTCTTCGCGCATATAGATCTATATGCTCTATGG 360  
QY 1409 CTTTCATGATGCTGCTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1468  
Db 361 CTTTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 1469 CACATATTTTCTACTAAATGCAGAGATTACCGGTGGCAATGGACAAGTTTCTCTCTGC 1528  
Db 421 CACATATTTTCTACTAAATGCAGAGATTACCGGTGGCAATGGACAAGTTTCTCTCTGC 479  
QY 1529 TGCATCAACTGCAATCTATGTTTACATGATATTCCTTTTACTACTATTTTTCAAAAACAA 1588  
Db 480 TGCATCAACTGCAATCTATGTTTACATGATATTCCTTTTACTACTATTTTTCAAAAACAA 538  
QY 1589 GATGATGCTTATTTCAACATCATTTTACTTTTGGATATATGGCGGTATTTAGCACAGC 1648  
Db 539 GATGATAGG-TTATTTCAACATCATTTTACTTTTGGATATATGGCGGTATTTAGCACAG- 596  
QY 1649 CTTGGGGATAATGTGTGGAGCGATTGGTTTACATGGGAACAAGTGCCTTTGTCGAAAAAAT 1708  
Db 597 CTTGGGGATAATGTGTGGAGCGATTGGTTTACCTGGGAACA---GTGCTTTGTGCAAAAAA 653  
QY 1709 CTATACATATGCAAAATTCAGTAGAGACCCCAAGAAACCTGGAACCTTTGGATCAATTTTC 1768  
Db 654 TTATCTAATGTGAAA--TTGATAGAGAACCAAGAAACCTGGGAACCTTGGATTCATTTC 711

QY 1769 TTTTTCATAGGGGTGGAACCTTG 1790  
Db 712 -TTTTCATACGGGTGGAACCTTG 732

Search completed: January 20, 2003, 16:29:38  
Job time : 2344 secs









Db 1046 GCGGCAAAACTGGATAAAAGCTATGATCCCTTACTGCATCCCTTTTCCGCTTCTTGCT 1105  
Qy 1075 GTGGCACTGCTTCTTCATCAATTTTCATAGCCATTTATTACCATGCTTCAAGAGCCATTC 1134  
Db 1106 TTTTCGATTGGATTGGTGTGCAACACCATGCTATCTTACCGTTTCACTGGCAGCCATAC 1165  
Qy 1135 CTTTGTGGAACAATGGTGGCGGTTTGTGTGCATCTGTTTTTTTGTATCTTCCCTCAANTC 1194  
Db 1166 CATTTGGCACAATGGTGTGCATATTTGCTCTGTGGCTTTTCATCTCTTTTCCCTTGGTTC 1225  
Qy 1195 TTTGTTGGTACAATACTTGGCCGGAATCTGTCAGTTCAGGCCAACTTTCTTGTGCTGTCA 1254  
Db 1226 TATTGGGAAGTGTAGTGTGTGAGAAATTTGGAGTGGCTCTCCCAACAATCCCTGTCGTGTTA 1285  
Qy 1255 ATGCTGTGCTCTGCTATACCGCGAGAAAATGGTTTCATGGAGCGCTGGGTTTATTGTTT 1314  
Db 1286 AGACAATTCACGCTCTATTTCTTGAGAAGAGTGGTATCTTACACCTCTGTTATCTCAT 1345  
Qy 1315 GCCTGGTGGAAATTTACCTTTTGGTTCAATCTTTATTGAAATGTAATTCATCTTCACGT 1374  
Db 1346 TGATGGTGGACTACTCCCTTCGGCAGCATCTTCATCGAGATGTACTTTGTGTTTACTT 1405  
Qy 1375 CTTTCTGGGCATATAGATCTATTATGCTATGCTATGGCTTCATGATGCTGGTGGTTATCC 1434  
Db 1406 CATTTCTGGAATCAAGAGTTTATTATGATATGTTTTCATGCTGCTGGTTTGTGTTATCC 1465  
Qy 1435 TGTGCAATGTGACTGCTGTGTGACTATTGTGTCACATATTTCTTACTAAATGCGAGAAG 1494  
Db 1466 TCATAATAGTCACCATATGTCACATATTGTGGGTACTTATTCTTGTGAATGCTGAGA 1525  
Qy 1495 ATTACCGGTGGCAATGGCAAGTTTCTCTCTGCTGCATCACTCAATCTATGTTTACA 1554  
Db 1526 ATTACCATTTGGCAATGSACATCATCTTCTCCGCTGCATCTACTGCTTGTATGCTACC 1585  
Qy 1555 TGATTTCTTTACTACTATTTTTCAAACAAAGATGATGCTGTTATTTCAAACATCAT 1614  
Db 1586 TATATCAATATACTATTATACGTGAAGACAAAGATGCTGGCTTCTTCCAGACAAGCT 1645  
Qy 1615 TTTTACTTTGGATATATGGGGTATTATAGACAGCCTTGGGATAAATGCTGGAGCGATTG 1674  
Db 1646 TCTACTTTGGCTACACTTGATGTTCTGCTCGCACTAGGAATCTTTGGGTTACTGTTG 1705

RESULT 4

US-09-887-576-812  
; Sequence 812, Application US/09887576  
; Patent No. US20020144047A1

GENERAL INFORMATION:

; APPLICANT: Budworth, P.  
; APPLICANT: Brown, D.  
; APPLICANT: Chang, H.  
; APPLICANT: Zhu, T.  
; APPLICANT: Han, B.  
; APPLICANT: Wang, X.  
; APPLICANT: Cooper, Bret  
; TITLE OF INVENTION: Promoters for regulation of plant expression  
; FILE REFERENCE: 1360, 001US1  
; CURRENT APPLICATION NUMBER: US/09/887,576  
; PRIOR FILING DATE: 2001-06-25  
; PRIOR APPLICATION NUMBER: US 60/213,848  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/214,087  
; PRIOR FILING DATE: 2000-06-23  
; PRIOR APPLICATION NUMBER: US 60/258,692  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 875  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 812

; LENGTH: 1803  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
US-09-887-576-812

Query Match 28.1%; Score 506; DB 10; Length 1803;  
Best Local Similarity 58.1%; Pred. No. 2, 4e-111;  
Matches 1006; Conservative 3; Mismatches 678; Indels 45; Gaps 5;  
Qy 29 GCTGCCCCGGAGCCGGGCGAGCAGCAGCAACACAGCTATCAAGATAAAGAGAAAGTTGT 88  
Db 90 GCGSGSCCGCGVTCGCCCTCCGAGTCAGACACCAAGTACAAAGTTGAAGAACCAAGTTAA 149  
Qy 89 CTTATCGATGAATACTGTTGGGCCCTTACCATAATCTCTCAAGAAACATATAAGTACTTTTC 148  
Db 150 GCTCTGGGTGAACAAGTCGGCCCTTACAATATTCACAAGAAACCTTACAACATATCACAG 209  
Qy 149 ACTTCCATTTCTGTGGGGTCAAAAAAAGTATCAGTCAATTACCATGAAACCTCTGGGAGA 208  
Db 210 CCTTCCATTTTGTCAACCACTCTGAGAACCT-- -GCGCATAAATGGGTGGTCTTGGAGA 266  
Qy 209 AGCACTTCAAGGGTTGAATTTAGTGGTCTGGGATATTAATTTTAAAGATGATGT 268  
Db 267 GGTCTCTGGGTGGAATAGAGCTGATTGATAGTCAGATTGACATAAAGTTCTTTAAGAAATGA 326  
Qy 269 GATGCCAGCCACTTACTGTGAAATTTAGATAAAGAAAGAGAGATGCCATTTGTATA 328  
Db 327 GGAGAGGGCTCCATTTGTACACTTGGAGCTTGATTTAAAGAGTTTCAGAGTTCTCCGA 386  
Qy 329 TGCCATTAATAAATCAATTACTTGTGTCAGAGATGTACATAGATGATTTACCAATATGGGTAT 388  
Db 387 TGCCATTTGATACTCATACTGTTTCAGTTTTCATGG-----GTTTTGT 431  
Qy 389 TGTGTTGAGGCTGATGAAATGGGAGAGATTACTATCTTTTGGACCTATAAANAAGTTGA 448  
Db 432 TGGGAGAGCGGACAAAAACAATGAAACAAGCGTTATCTCTACACACACAAGAGCATCT 491  
Qy 449 AATAGTTTTAATGGAATCGAATTTGATGTTTATCTTAATCTAAGTCTAGTGAAGGAAGTGAA 508  
Db 492 TGTCAAAATACAATGATAACAGGATCATTCATGTTAATCTACCCCAAGAGTCTCTAAGCT 551  
Qy 509 ACTGGTTCCAAATFACTAAATCCAGATGTCATATTTCAGTAAATGGAAAGAGTCAGATGT 568  
Db 552 TCTTGAAGCTGGTGAAGAAATGGACATGACATATTCAGTGAAGTGGTTACAACAGATCT 611  
Qy 569 GAAATTTGAAGATCGAATTTGACAAATATCTTGAATCGGTCCTTTTTCACACATCGGATCA 628  
Db 612 GACCTTTGCACCGCGCTTTTGAAGTATACCTGGATTATCCTTTTTCAGCACACAGATTCA 671  
Qy 629 TTTGTTTTCAATTTTCAACTCTTCATGATGTTGATGTTCTTGGTGGGCTTAGTTTCAAT 688  
Db 672 TTTGTTCTCCATTTTCAATTTCTTCATGATGTTTATTTTCCCTCACTGGTTTGGTTTCAAT 731  
Qy 689 GATTTTATGAGAACATTAAGAAAGATTATGCTCGGTA---CAGTAAAGAGAGAAAT 745  
Db 732 GATATTGATGCGGACATTCAGAAATGACATATGCAAAAGTATGCTCGTGAAGATGATGATCT 791  
Qy 746 GGATGATATGGATAGAGACCTAGGAGATGAATATGATGGAACAGGTCATCGGAGATGT 805  
Db 792 GGAGTCACCTGGAGAGAGATGTTAATGAGGAATCTGGATGGAAGCTTGTTCATGTTGATGT 851  
Qy 806 ATTTAGACCATCAAGTCACCCACTGATATTTTCCCTCTCTGATTGGTTCGATGTCAGAT 865  
Db 852 ATTTGCCCTCTCGAAGTTTGGCCCTTTCTTTGTCAGTTTGGCATCGGCACTCAGCT 911  
Qy 866 ATTTGCTGTGCTCTCATCCTTATTATTGTTGCAATGATAGAGATTTTATATCTG--- 921  
Db 912 GGCAGCTCTTATCCTTCTTGTATTGCTATTGGCCATTTGTTGGCATGCTATATGTTGGCCG 971  
Qy 922 -----AGAGGGGATCAATGCTCAGTACAGCCATATTTGCTCTATGCTGCTAGC 968  
Db 972 GCGATATTTGTCAGCGAGGGTCTATCATCAACAACCTTCATTCTGCTGCTATGCTCTTACA 1031  
Qy 969 TCTCCAGTGAATGGTTATTGTCAGAGAACTCTGTATGCTAGACAAG-----AGGAAG 1021  
Db 1032 TCTTTTATTTCTGGCTATGTCAGTGKGGCCCTTACTCAAGGAATGTTGGTATGGTTAA 1091  
Qy 1022 GAGATCGATAAAGCAGATGTTTATTGGGGCATTCCTTATCCAGCTATGTTGTTGGCAC 1081

Db	1092	AACTGGATAAAGCGTATGATCTTACAGCATCCCTATTTCATCTTGCTGTGCAAT	1151
Qy	1082	TGCTCTTCATCAANTTTCATAGCCATTTATTTACCATGCTTCAAGAGCAATTCCTTTTGG	1141
Db	1152	TGGCTTTGTTCTGAATACAATTCGTATCTTCAACGATCATTTAGCGCAATACCAATTCGG	1211
Qy	1142	AACAATGGTGGCGGTTGTTGGCATCTGTTGTTTTTTTGTCTTATCTTCTCTAAATCTGTGTGG	1201
Db	1212	CACAATGGTTGTCTATGTTGGTCTTGGGCTTCATTTCTTTTCGGCTGTGCTCTCTGGG	1271
Qy	1202	TACAATACTTGGCGGAATCTGTCAGTGCAGCCCAACTTCTCTGTCGTGCAATGCTGT	1261
Db	1272	AACTGTTGGTAGAAGTGGTGGTCTTCAACAAATCCCTGCCGAGTAAGAAGATAT	1331
Qy	1262	GCCTGCTCTATACCGGAGAAAAATGGTTTCATGAGGCTCGGGTTATTTGTTGCCCTGGG	1321
Db	1332	TCCAGCGCTATTCCAGAGAAGTGGTACCTTACACCTTCTGTTATCTCATTTGATGG	1391
Qy	1322	TGCAATTTACCTTTTGGTTCAATCTTTATTTGAATGTATTTTCATCTTCAGTCTTTCTG	1381
Db	1392	TGGGCTGCTTCCCTTTGGCAGTATCTTCAATTGAGATGTACTTTGTGTTACGTCATTTCTG	1451
Qy	1382	GGCATATAAGATCTATTATGTCATGGCTTCATGATGCTGGTGTGTTATCTGTGTCAT	1441
Db	1452	GACTACAGGTTTATTATGCTACGGCTTCATGCTGGCTTTGTTCATCTCTCTAAT	1511
Qy	1442	TGTGACTGTCTGTGACTATTGCTGTGCACATATTTTCTACTAAATGCAGAAAGATACCG	1501
Db	1512	AGTTACAATCTGTGTACATATTGTGGGTACTTATTTCTGCGAATGCTGAGAACTATCA	1571
Qy	1502	GTGGCAATGGACAAGTTTCTCTCTGTGTGATCAACTGCAATCTATGTTTACATGTATTC	1561
Db	1572	CTGGCAATGGACTTCAATTTCTCTCTGCTGCATCAACTGCGTTATATGTATCTATATTC	1631
Qy	1562	CTTTTACTACTATTTTTCAAAACAAGATGTAGGCTTATTTTCAACATCATTTTACTTT	1621
Db	1632	CATCTACTATCATGTAAAGACAAGAATGTAGGTTTTCAGACGAGTTTCTACTTT	1691
Qy	1622	TGGATATATGGCGGTATTTTAGCACACGCTTGGGATAATGTGTGGAGCGATTGGTTACAT	1681
Db	1692	CGGTTACACATTGATCTCTGCTTGGCTAGGCATACTTTTGTGGTGTATTGGCTATCT	1751
Qy	1682	GGGAACAAGTGCCTTTGTCGGAAAAATCTATACTTAATGTGAAAATGTGACTAG	1733
Db	1752	AGGGTCAACTCTTTTGTGAGAGAAATCTACAGAACAATCAATATGTCATTAG	1803

## RESULT 5

```

US-09/887-576-794
; Sequence 794, Application US/09887576
; Patent No. US2002144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 794

```

```

QY 943 CAGCCATATTTCTCTATGCTGCTACGCTCCAGTGAATGGTATTTTGGAGGAAGTCTGT 1002
D 995 CATTCAATTTGTTGTTGCTTACTGCTTCTCATCTCGATATCTCAGTGTGTCACATTT 1054
QY 1003 ATGCTAGACAGGAGGAGAGATGATTAAGCAGATGTTTATGCGGCAATCTCTATCC 1062
D 1055 ATTCACGGCATGGGGGAAACTGATCAAGCAATGATTATGACAGCATCATTATTC 1114
QY 1063 CAGCTATGCTGCTGCTCTCTCTCATCAATTTTCATAGCCATTTATACCATGCTT 1122
D 1115 CTTTATGCTTTTGGAAATGGCTTAGTCTTAACACAACTGCTATATCTATCATCAT 1174
QY 1123 CAAGACCATTCCTTTTGGAAATGGCTTAGTCTTAACACAACTGCTATATCTATCAT 1182
D 1175 TAGCTGCCATACCATTTGCTACTATGCTGCTTCTCATCTGCTGGCTTCATATCTT 1234
QY 1183 TTCTCTTAATCTTTGCTGCTGCTCTCTCATCAATTTTCATAGCCATTTATACCATGCTT 1242
D 1235 TCCCTCTGCTCTTTTGGAACTGTTGTTGGTAGAACTGGAGTGGTCCGCCAAATAATC 1294
QY 1243 CTGCTGCTCAATCTGCTGCTCTCTCATCAATTTTCATAGCCATTTATACCATGCTT 1302
D 1295 CATGAGAGTAAAGACTATCTCTGCCCTATCCCTGAGAAAGATGGTACCTTCAGCCCTT 1354
QY 1303 CGGTATTTGCTGCTGCTGCTGCTCTCTCATCAATTTTCATAGCCATTTATACCATGCTT 1362
D 1355 CTGCTATTTGCTGCTGCTGCTGCTCTCTCATCAATTTTCATAGCCATTTATACCATGCTT 1414
QY 1363 TCATCTTCAGCTCTTTTGGGCAATGAGATCTATATGCTATGCTGCTGCTGCTGCTGCT 1422
D 1415 TTGCTCTCACATCATTTTGGAACTACAAGGTGATGATGATGATGATGATGATGATGAT 1474
QY 1423 TGCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482
D 1475 TCTTTTGTATCTCTATATGTCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1534
QY 1483 TAAATGCAAGATTTACCGTGGCAATGACAAAGTTTCTCTCTGCTGCTGCTGCTGCTGCT 1542
D 1535 TGAATGCGGAGAACTACCACATGCGAGTGGCTTCTCTCTGCTGCTGCTGCTGCTGCTGCT 1594
QY 1543 TCTATGTTTACATGATTTCTCTTTTACTACTATTTTCTCAAAACAAAGATGATGCTTAT 1602
D 1595 TCTATGTTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1654
QY 1603 TTCAACATCATTTTACTTTTGGATATGCGGTATTTAGCACAGCTTGGGGATATGT 1662
D 1655 TTCAGACAAGTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1714
QY 1663 GTGGAGGATTTGGTTTACATGGGAA 1686
D 1715 GCGGTAAGCTCTTCTACACCGGTA 1738

```

```

RESULT 6
US-10-201-964-3
: Sequence 3, Application US/10201964
: Publication No. US2003008356A1
: GENERAL INFORMATION:
: APPLICANT: Ono Pharmaceutical Co., Ltd.
: TITLE OF INVENTION: A NOVEL POLYPEPTIDE, A CDNA ENCODING THE SAME, AND USE
: FILE OF INVENTION: OF IT
: FILE REFERENCE: Q62911
: CURRENT APPLICATION NUMBER: US/10/201,964
: CURRENT FILING DATE: 2002-07-25
: PRIOR APPLICATION NUMBER: US/07/62,467
: PRIOR FILING DATE: 2001-02-07
: PRIOR APPLICATION NUMBER: JP 10-224308
: PRIOR FILING DATE: 1998-08-07
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 3
: LENGTH: 2083
: TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Clone OAH047 - human umbilical vein HUV-EC-C
; OTHER INFORMATION: endothelial cell
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (19)..(1944)
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: (19)..(87)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (88)..( )
US-10-201-964-3

Query Match      12.6%; Score 227.4; DB 9; Length 2083;
Best Local Similarity 52.3%; Pred. No. 8.4e-45;
Matches 635; Conservative 0; Mismatches 556; Indels 24; Gaps 5;

QY 520 ATACTAAATCCAGATGTCATATTCAGTAAATGCAAAAGTCAGATGTAATTTGAAG 579
D 752 AGAATCAGCTGTACTTCACTTCTGTCACCTGGGAGAAAGTATCAATGGGCT 811
QY 580 ATCGATTTGACAAATATCTTGTATCGTCTCTTTTCAACATCGGATTCATTTGGTTTCAA 639
D 812 CTCGCTGGGACATTTACCT-----GACCATGAGTGACGTCAGATCCACATGTTTCTA 865
QY 640 TTTTCAACTCTTCATGATGATGATCTTCTTGTGGGGCTTAGTTTCAATGATTTTAATGA 699
D 866 TCATTAACCTCGTGTGTTGTTGCTTCTTCTCTGAGGATCTCTGAGCATGATTAATCATTC 925
QY 700 GAACATTAAGAAAGATTTATCTCGTACAGTAAAGAGAGAAAGATGATGATGATA 759
D 926 GGACCTCCGAGAGACATTTGCCAACTACAAGAGGAGTACATTTGA----- 974
QY 760 GAGACTAGGAGATGAATATGATGGAACAGATGTCATGAGATGTTATTAGACCATCAA 819
D 975 -AGACACCATGAGGAGTCTGGTGGAAAGTTGGTGCACGGCAGCTCTTCAGGCCCCCC 1033
QY 820 GTCACCCACTGATATTTTCTCTCTGATGTTGTTGTTGAGATGTCAGATATTTGCTGTGTC 879
D 1034 AGTACCCCATGATCTCAGCTCCCTGCTGGGCTCAGGCATTCAGCTGTCTGTATGATCC 1093
QY 880 TCATCGTTATTTATGTTGCAATGATAGAAGATTTATATACATGAGA--GGGATCAATGC 936
D 1094 TCATCGTATCTTTGTAGCCATGCTTGGGATGCTGTCGCCCTCCAGCCGGGAGGCTCTCA 1153
QY 937 TCAGTACAGCCATATTTTGTCTATGCTGCTACGTCCTCCAGTGAATGTTATTTTGGAGAA 996
D 1154 TGACCACAGCTGCTTCTCTCATGTTGTTGCGGATTTTCTGCTGCC 1213
QY 997 GTCGTATGCTAGACAGGAGGAGATGATTAAGCAGATGTTTATTTGGGCAATTC 1056
D 1214 GTCGTACCGCACTTTAAAGGCCATCGTGGTGAAGAGAGACCTTCTGTACGGCACTC 1273
QY 1057 TTATCCAGCTATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1116
D 1274 TGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1333
QY 1117 ATGCTTCAAGACCATTCCTTTTGGAAACATGTTGGGCTTTGTTGATCTGTTTGTG 1176
D 1334 ACTCATCAGGAGCGTGGCTTTTCCACCATGTTGGTGGCTCTGCTGTCATGTCGTCGG 1392
QY 1177 TTATCTCTCTTAATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1236
D 1393 ATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1450
QY 1237 ACTTTCCTTGTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1296
D 1451 ACAACCTGCTGCGACCAACAGATTTCCCGGACAGATCCCGGACAGCGGGTGGTACATGA 1510
QY 1297 AGCCTGCGGTTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1356

```

Db	1511	ACCGATTGTGGGCACTCCTCATGGCTGGGATCTTGCCCTTCGGCGCCATGTTTCATCGAGC	1570
QY	1357	TGATATTCATCTTCACGCTCTTCTGGGCATATAGATCTATATGTCATGGCTTCATGA	1416
Db	1571	TCTTCTTCATCTTCAGTGCTATCTGGGAGATCAGTTCTATACCTCTTTGGCTTCCTGT	1630
QY	1417	TGCTGGGCTGGTATCCTCTGTCATTTGACTGTCTGTGACTATTTGTGTGCACATATT	1476
Db	1631	TCCTTGTGTTTCATCATCCTCGTGTATCCGTTTCACAAATCAGCATGTCATGGTGTACT	1690
QY	1477	TTCCTACTAAATGAGAGAATACCGGTGGCAATGGAGCAAGTTTCTCTGCTGTCATCAA	1536
Db	1691	TCCAGCTGTGCGAGAGATATCCGCTGGTGGTGAGAAAATTCCTAGTCTCCGGGGCT	1750
QY	1537	CTGCAATCTATGTTTACATGATTTTCCCTTTTACTACTATTTTTCAAAACAAAGATGATG	1596
Db	1751	CTGCAATCTACGTCCTGGTGTGATCCATCTTTTATTTTCGTTAAACAGCTGGACATCTGG	1810
QY	1597	GCTTATTTCAAACATCATTTTACTTTTGGATATATGGCGGTATTTAGCACAGCCTTGGGGA	1656
Db	1811	AGTTTCATCCCTCTCTCTACTTTGGCTACACGGCCCTCATGGTCTTGTCTCTCTGGC	1870
QY	1657	TAATGTGTGGAGCGATTGGTTTACATGGGAACAAGTGCCTTTTGTCCGAAAAATCTATACTA	1716
Db	1871	TGCTAACGGGTACCATCGGCTTCTATGCAGCCTACATGTTTGTTCGCAAGATCTATGCTG	1930
QY	1717	ATGTGAATAATTGACT	1731
Db	1931	CTGTGAAGATAGACT	1945
RESULT 7			
US-10-201-964-2			
; Sequence 2, Application US/10201964			
; Publication No. US20030008356A1			
; GENERAL INFORMATION:			
; APPLICANT: Ono Pharmaceutical Co., Ltd.			
; TITLE OF INVENTION: A NOVEL POLYPEPTIDE, A CDNA ENCODING THE SAME, AND USE			
; TITLE OF INVENTION: OF IT			
; FILE REFERENCE: 062911			
; CURRENT APPLICATION NUMBER: US/10/201,964			
; CURRENT FILING DATE: 2002-07-25			
; PRIOR APPLICATION NUMBER: US/09/762,467			
; PRIOR FILING DATE: 2001-02-07			
; PRIOR APPLICATION NUMBER: JP 10-224308			
; PRIOR FILING DATE: 1998-08-07			
; NUMBER OF SEQ ID NOS: 5			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 2			
; LENGTH: 1926			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-201-964-2			
Query Match 12.6%; Score 226.4; DB 9; Length 1926;			
Best Local Similarity 52.2%; Pred. No. 1.4e-44;			
Matches 634; Conservative 0; Mismatches 556; Indels 24; Gaps 5;			
QY	520	ATACTAAATCCAGATGTCATATTCAGTAAATGGAAAAAGTCAGATGTGAAATTTGAAG	579
Db	734	AGAATCAGCTGTACTTCCACTACTCTGTCCACTGGGAGGAAGTGATATCAAAATGGSCCT	793
QY	580	ATCGATTTGACAAATCTTGATCCGTCCTTTTTCACATCGGATTCATTTGGTTTCAA	639
Db	794	CTCGCTGGGACACTTACCT-----GACCATGAGTGACGTCCAGATCCACTGGTTTCTA	847
QY	640	TTTTCAACTCCTTCATGATGGTATCTTCTGTGGGCTTGTGTTCAATGATTTTAATGA	699
Db	848	TCATTAACCTCGTGTGTGCTCTCTCTCTGTCAGGTATCCTCGAGCATGATATCATTC	907
QY	700	GACATTTAAGAAAGATTATGCTGGGTACAGTAAAGAGGAAGAAATGGATATGGATA	759

Db	908	GGACCCCTCCGAAGGACATTTGCCAACTACAAAGAGGAGATGACATTGA-----	956
QY	760	GAGACCTAGGAGATGAATATGGATGGAAACAGGTGCATGGAGATGTATTTAGACCATCAA	819
Db	957	-AGACACCATGGAGAGCTGGGTGGAAAGTTGGTGACGGCGAGCTTTCAGGCCCCCC	1015
QY	820	GTCACCCACATGATATTTTCTCTCTGATTGGTTCTGGATGTCAGATATATTTGCTGTGTCTC	879
Db	1016	AGTACCCCATGATCCTCAGCTCCCTGCTGGCTCAGGCATTCAGCTGTTCTGTATGATCC	1075
QY	880	TCATCGTTATATTTGTTGCAATGATAGAAGATTTATATACTAGAGA---GGGGATCAATGC	936
Db	1076	TCATCGTCATCTTTGTAGCCATGCTTGGGATGCTGCCCTCCAGCGGGAGCTCTCA	1135
QY	937	TCAGTACAGCCATATTTGCTATGCTAGCTCCAGTCTCCAGTGAATGGTTATTTGGAGGAA	996
Db	1136	TGACACAGCCCTGCTTCCTCTTCATGTTTATGGGGGTGTTTGGCGGATTTCTGCTGGCC	1195
QY	997	GTCTGTATGTAGACAGGAGGAGATGGATAAAGCAGATCTTTATTTGGGGCATTTCC	1056
Db	1196	GTCTGTACCCCACTTTTAAAGGCCATCGGTGGAAGAAGGAGCCTCTGTACGGCACTC	1255
QY	1057	TTATCCCAAGTATGGGTGTGGCACTGCCCTTCTTCATCAATTTTCATAGCCATTTATTACC	1116
Db	1256	TGTACCCCTGTGTGGTGTGGTTCGTTTCGTTATTTGAATTCGTTTCATTTGGGGAAGC	1315
QY	1117	ATGCTTCAAGAGCCATTCCTTTTGGACAATGGTGCCGTTTGTGTCATCTGTTTGTG	1176
Db	1316	ACTCATCAGGAGCGGTGCCCTTTCCCACTAGTGGTGGCTCTGCTGGCATGTG-GTTCGGG	1374
QY	1177	TTATTTCTCCTCTAAATCTTTGGTACAACTACTTGGCCGAAATCTGTCAGGTCAGGCCA	1236
Db	1375	ATCTCCCTGCCCTCGTCTACTTGGGCTACTACTTGGCTTCCGAAAGCAG--CCATATG	1432
QY	1237	ACTTTCCTTCTGTGTCATATGCTGTGCTCTATATACCGGAGAAAAATGTTTCATGG	1296
Db	1433	ACAACCCCTGTGCGCACCAACAGATTTCCCGCGCAGATCCCGAGCAGCGGTGTACATGA	1492
QY	1297	AGCTCGGGTATTTGTTGCTGGGTGGAATTTACCTTTTGGTTCAATCTTTATTTGAAA	1356
Db	1493	ACCGATTTGGGGCATCCTCATGCTGGGATCTTGCCCTTCGGGCGCATGTTTCATCGAGC	1552
QY	1357	TGATTTTCATCTTCACGCTCTTCTGGGCATATAAGATCTATATGTCATGGCTTCATGA	1416
Db	1553	TCCTTCTCATCTTCAGTGCTATCTGGGAGAAATCAGTCTCTATACCTCTTTGGCTTCCTGT	1612
QY	1417	TGCTGGTGTGTTATCCTCTGTCATTTGCTGCTGTGCTGTGACTATTTGTGTGCACATATT	1476
Db	1613	TCCTTGTGTTTCATCATCTCTGCTGCTGCTGTTTTCACAAATCAGCATGCTGCTGCTACT	1672
QY	1477	TTCTACTAAATGCAGAGATACCGGTGGCAATGGACAAGTTTCTCTCTGCTGCATCAA	1536
Db	1673	TCCAGCTGTGTGCAGAGGATTCACCGGTGGTGGAGAAAATTTCCCTAGTCTCCGGGGGCT	1732
QY	1537	CTGCAATCTATGTTTACATGATTTCTCTTTTACTACTATTTTTCAAAACAAAGATGTATG	1596
Db	1733	CTGCAATCTACGCTCGTGTATCCCATCTTTATTTTCGTTAAACAACCTGCACATCGTGG	1792
QY	1597	GCTTATTTCAACATCATTTTACTTTTGGATATATGGCGGTATTTAGCACAGCCTTGGGGA	1656
Db	1793	AGTTTACCCCTCTCTCTCTACTTTTGGCTACACGCGCCCTCATGGTCTTGTCTTCTGCG	1852
QY	1657	TAATGTGTGAGCGATTTGTTTACATGGGAACAGTGCCTTTTGTCCGAAAAATCTATACTA	1716
Db	1853	TGCTAACGGGTACCATCGGCTTCTATGCAGCCTTACATGTTTGTTCGCAAGATCTATGCTG	1912
QY	1717	ATGTGAATAATTGAC	1730
Db	1913	CTGTGAAGATAGAC	1926

; Sequence 754, Application US/09770791  
; Patent No. US20020062014A1  
; GENERAL INFORMATION:  
; APPLICANT: Gorlach, Jorn  
; APPLICANT: An, Yong-Qiang  
; APPLICANT: Hamilton, Carol M.  
; APPLICANT: Price, Jennifer L.  
; APPLICANT: Raines, Tracy M.  
; APPLICANT: Yu, Yang  
; APPLICANT: Rameaka, Joshua G.  
; APPLICANT: Page, Amy  
; APPLICANT: Matthew, Abraham V.  
; APPLICANT: Ledford, Brooke L.  
; APPLICANT: Woessner, Jeffrey P.  
; APPLICANT: Haas, William David  
; APPLICANT: Garcia, Carlos A.  
; APPLICANT: Kricker, Meja  
; APPLICANT: Slader, Ted  
; APPLICANT: Davis, Keith R.  
; APPLICANT: Allen, Keith  
; APPLICANT: Hoffman, Neil  
; APPLICANT: Hurban, Patrick  
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis  
; FILE REFERENCE: 2029 (PARA-018PRV)  
; CURRENT APPLICATION NUMBER: US/09/770,791  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 60/178,480  
; PRIOR FILING DATE: 2000-01-27  
; NUMBER OF SEQ ID NOS: 999  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 754  
; LENGTH: 347  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-770-791-754

Query Match 9.6%; Score 172.6; DB 10; Length 347;  
Best Local Similarity 69.3%; Pred. No. 4.7e-32;  
Matches 235; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 1285 AATGGTTCATGAGCGTCGCGTATTTGGCTGGGTGGAATTTACCTTTTGGTCAA 1344  
DB 9 AATGGTACCTTCACTCCATCAGTCGCTCCCTGATGGGGGTTTGTACCTTTTGAAGCA 68  
QY 1345 TCCTTATTAAGATTTTCACTTTCACCTCTTCTGGGCATATAGATCTATTATGCT 1404  
DB 69 TCTTATTCAGATGTACTTTGTCTTCACATCCTTCTGGGAATTACAAGGCTTACTATGTGT 128  
QY 1405 ATGGCTTCATGATGCTGGTGTATCCTGTGATTTGTGACTGTCTGTGACTATTATG 1464  
DB 129 ATGGATTCATGTTACTGGTTTGTGATCTCGTCATAGTACGGTGTCTGTGACATCG 188  
QY 1465 TGTGCACATATTTTCTACTAAATGCAAGATTTACGGTGGCAATGGACAAGTTTCTCT 1524  
DB 189 TGGGTACATATTTCTGCTGAAATGCAAGAACTATCAGTGGCAGTGGACTTCATTTTCT 248  
QY 1525 CTGCTGCATCAACTCAACTATGTTTACATGATATTCCTTTTACTACTATTTTCAAAA 1584  
DB 249 CTGCTGCTCGACGGGTCTATGTGTACTTATCTACTTACTTACTACTACTACGTAAGA 308  
QY 1585 CAAAGATGTATGGCTTATTTCAAACTATTTTACTTTG 1623  
DB 309 CCAAGATGTCGGATTTCTCCAGACAAGCTTCTACTTTG 347

RESULT 9  
US-09-938-842A-798  
; Sequence 798, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SAME, AND METHODS OF USE  
; FILE REFERENCE: SRIPI300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 798  
; LENGTH: 1980  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-798

Query Match 9.5%; Score 171.8; DB 9; Length 1980;  
Best Local Similarity 49.2%; Pred. No. 1.5e-31;  
Matches 551; Conservative 0; Mismatches 552; Indels 18; Gaps 3;

QY 614 TCAACATCGGATTCATTGGTTTCAATTTTCAACTCCTTCATGATGGTGATCTCTTGTGT 673  
DB 873 TGATGATCAGATTCATTGGTTCTCAATCGTTAATCTATGATGATGTTCTTCTCTCTC 932  
QY 674 GGGCTTAGTTTCAATGATTTTAAATGAGACATTAAGAAAAGATTATGCTCGGTACAGTAA 733  
DB 933 TGGTATGGTCGTATGATCATGCTAAGAACACTCTACCGAGATATTTCCAACATCAATCA 992  
QY 734 AGAGNAGAAATGGATGATATGATAGACACCTAGGAGATGAATATGATGGAACACGCT 793  
DB 993 -----GTTAGAGAGTCATGAAGAGGCCCTCGAAGAGACGGGTGGAAATTTGGT 1040  
QY 794 GCATGGAGATGATTTAGACCATCAAGTCACCCACTGATATTTCTCTCTCTGATGGTTC 853  
DB 1041 ACACGGAGATGTTTTCAGACCCCAACAAACCGGAGTTGCTTTGTGTTATGACAGGCAC 1100  
QY 854 TGGATGTCAGATATTTGCTGCTCTCATCGTTATTTATGTTGCAATGATAGAGATTT 913  
DB 1101 TGGAGTTCAATGCTTTGGAATGATTTTGTCCACCATGATCTTTGCTTGGCTTTTGGT 1160  
QY 914 ATATCTGAGAGGGGATCAATGCTCAGTACAGCCATA---TTTGTCTATGCTGCTACGTC 970  
DB 1161 ATCTCCTCAAAACCGTGGTGGTCTTATGACGGCTATGCTTTTGTGTTGGTTCATGGG 1220  
QY 971 TCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTAGACAAGGAGGAGGATGGAAT 1030  
DB 1221 ACTATTGGCAGGATACGCGTCTTCACGCGCTCTACAAGACGTTTAAGAGGAACCGAATGGA 1280  
QY 1031 AAGCAGATGTTTATTTGGGGGATTCCTTATCCAGCTATGCTGCTGCGACATGCGCTCTCT 1090  
DB 1281 AAGAACGCTCTGAAACCGCATTCATGTTTCCCGCTACCGCTTTGTGCGCTCTCTCTGT 1340  
QY 1091 CATCAATTTTATGACATTTTATTTACCATGCTTCAAGAGCATTCCTTTTGGAAACATGCT 1150  
DB 1341 CCTTATGGATTAATCTGGGGACAAAATATCCCGGTGACGTCCTCATTTGGTACAAATGTT 1400  
QY 1151 GGCGGTTTGTTCATCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 1210  
DB 1401 TGCTTTGGTTGCTCTGGTTTGGTATCTCTGTTCCACTGTTTTCATCGGTGGGTACAT 1460  
QY 1211 TGGCGAAATCTGTCAGGTACCGCCCACTTCTTGTCTGTCGATGCTGCTGCTGCTGCTG 1270  
DB 1461 TGGTTTCAGAAAACCTGCACCGGAAGA---TCCAGTGAACCAACACAGATCCCGGCA 1517  
QY 1271 TATACCGGAGAAAATGTTTCTCATGAGCTCGGTTATTTGTTGCTGGGTGGAAATTT 1330  
DB 1518 GATCCCAACACAGGCTTGGTATACGAACCAATCTTCCCAATTTTGATTTGGAGGATCT 1577  
QY 1331 ACCTTTTGGTTCAATCTTTTATTGAAATGATTTTCTATCTTCTACGTCCTTTCTGGGCATATA 1390



DB 1578 CCACATTTGGCGCAGCTTTCATCGAGCTCTTCTCATACTCACTCAACTCAATATGCTACATCA 1637

QY 1391 GATCTATTATGTCTATGGCTTCATGATGCTGGTGCTGGTTATCCGTGCATTTGCTGACTGT 1450

DB 1638 GTTCTACTACATATTCGCGCTTCTCTCATGCTCTCATCATCTTTTGATCATCACTTTGGCG 1697

QY 1451 CTGTGTGACTATFTGTGTGCACATAATTTTTCTACTAAATCAGAAGATTAACCGTGGCAATG 1510

DB 1698 AGAGATCACGGTCGCTCTGTATTTTTCAGCTGTGTAGTGAAGACTATCAATGCTGGTG 1757

QY 1511 GACAAGTTTTTCTCTCTGCTGCACTAACCTGCAATCTATGTTTACATGTAATCTCTTTTACTA 1570

DB 1758 GAGATCTTTACATTAACATCAGGCTCTCCGCGGTTTTACCTCTCTCTTTACGCACTCTTTA 1817

QY 1571 CTATTTTTTCAAACAAGATGTATGGCTTATTTCAAACATCATTTTACTTTGGATATAT 1630

DB 1818 CTTCTACACCAAACTTGAGATCACAAAGCTGTCTCTGAGTGCTCTACTTTGGGTACAT 1877

QY 1631 GCGCGTATTTTAGCACAGCCTTTGGGATAATGTGTGGAGCGATTTGGTTACATGGGAACAAG 1690

DB 1878 GCTCATAGTTTCATACGATATCTTCGTCCTTACTGTCGCAATGGTTTTCTACGCCATGCTT 1937

QY 1691 TGCCTTTTCCGAAAACTTACTACTAAATGTAAGAAATGACT 1731

DB 1938 TTGGTTCACCAGGCTTATCTACTCTTCGTTAAAATCGATT 1798

RESULT 10

US-09-910-664-30/c

: Sequence 30, Application US/09910664

: Publication No. US20020194646A1

: GENERAL INFORMATION:

: APPLICANT: POGUE, Greg P.

: APPLICANT: DELLA-CIOPPA, Guy R.

: APPLICANT: WOLFE, Gerson M.

: APPLICANT: ZHENG, Wenjin

: TITLE OF INVENTION: METHODS OF CREATING DWARF PHENOTYPES IN

: TITLE OF INVENTION: PLANTS

: FILE REFERENCE: 00801018900US01

: CURRENT APPLICATION NUMBER: US/09/910.664

: CURRENT FILING DATE: 2001-07-20

: NUMBER OF SEQ ID NOS: 122

: SOFTWARE: FastSeq for Windows Version 4.0

: SEQ ID NO 30

: LENGTH: 1132

: TYPE: DNA

: ORGANISM: Arabidopsis thaliana

: US-09-910-664-30

	Query Match	9.4%	Score 170;	DB 9;	Length 1132;
	Best Local Similarity	50.0%;	Pred. No. 3.3e-31;		
	Matches 481;	Conservative 0;	Mismatches 475;	Indels 6;	Gaps 2;
Qy	782	ATGNAACAGGTGCATGGAGATGTRATTTAGACCATCAAGTCACCCACTGATATTTTCCTC	841		
Db	1132	ATGGAAGCTTGTACACGGAGATGCTTCAGGCCACCGGTGAACCTCGTTATTTATGTTGTGT	1073		
Qy	842	TCTGATTTGGTCTCGATGTCAGATATTTGCTGTCCTCATCGTTATTTATGTTGCAAT	901		
Db	1072	TTATGTTGGTACAGGTGTTGCAGATCTTCGGAATGTCATCTTACAATGATGTTTGGCGTT	1013		
Qy	902	GATACAAGATTTATATCTAGAGGGGATCAATGCTCAGTACAGCCATATTTGTC---	958		
Db	1012	GCTTGGCTTCTTATCTCCATCCACAGAGGAGGCTTATGACTGCCATGGTTCCTTTGTTG	953		
Qy	959	TGCTGCTACGPTCCAGTGAATGTTATTTTGGAGGAAGTCTGATGCTAGACAAGGAGG	1018		
Db	952	GTTTTTCATGGCATATTCGCTGGTACTCCTCGCTCGCCCTTCACAAAATGTTCAARAG	893		
Qy	1019	AAGAGATGGATAAAGCAGATGTTTATTTGGGCCATTCCTTATCCAGACTATGGTGTGTG	1078		
Db	892	AAACAAGTGAAGAGAATGACCTTGAAGACTGCAATCATGTTTCCCGGTATCTTTTTCG	833		

```

RESULT 11
US-09-887-576-471
; Sequence 471, Application US/09887576
; Patent No. US20020144047A1
;
; GENERAL INFORMATION:
;   APPLICANT: Budworth, P.
;   APPLICANT: Brown, D.
;   APPLICANT: Chang, H.
;   APPLICANT: Zhu, T.
;   APPLICANT: Han, B.
;   APPLICANT: Wang, X.
;   APPLICANT: Cooper, Bret
;   TITLE OF INVENTION: Promoters for regulation of plant expression
;   FILE REFERENCE: 1360, 001US1
;   CURRENT APPLICATION NUMBER: US/09/887,576
;   CURRENT FILING DATE: 2001-06-25
;   PRIOR APPLICATION NUMBER: US 60/213,848
;   PRIOR FILING DATE: 2000-06-23
;   PRIOR APPLICATION NUMBER: US 60/214,087
;   PRIOR FILING DATE: 2000-06-23
;   PRIOR APPLICATION NUMBER: US 60/258,692
;   PRIOR FILING DATE: 2000-12-29
;   NUMBER OF SEQ ID NOS: 875

```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 471
; LENGTH: 4300
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-471

Query Match      8.0%; Score 143.2; DB 10; Length 4300;
Best Local Similarity 64.5%; Pred. No. 1.4e-24;
Matches 214; Conservative 0; Mismatches 118; Indels 0; Gaps 0;

QY 1341 TCAATCTTTATTGAATGCTATTTTCATCTTCACGCTCTTTCTGGGCATATAGATCTATTAT 1400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3825 TAAATGTAGTGGCTATTTTGAAGTGTCTCTTTTGTTCGACAGGCTCTACTAT 3884

QY 1401 GTCTATGGCTTTCATGCTGGTGGTGTATCTCTGTCATTCGACTGCTGCTGTGACT 1460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3885 GTGTATGGATCAATGTTACTGGTTTTTGTGATTCCTGTCATAGTGAGGGTGTGTGACA 3944

QY 1461 ATTGTGTGCACATATTCTTACTAAATGCAAGATTACCGGTGGCAATGGACAAGTTTT 1520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3945 ATCGTGGGTACATATTCTCTGCTGAATGCAGAGAACTATCACTGGCAGTGGACTTTCATT 4004

QY 1521 CTCTGCTGCATCACTGCAATCTATGTTTACATGATTTCCCTTTTACTACTATTTTTC 1580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4005 TTCTCTGCTCTTCGACGGCTGTATGTACTTATCTACTCCATCTATTACTACTACGTA 4064

QY 1581 AAACAAAGATGATGCTGTTATTTCAACATCATTTTACTTTTGGATATATGGCGGTATT 1640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4065 AGACCAAGATGTCGGATTCCTCCAGACAAAGCTTCTACTTTGGATACACCATGATGTT 4124

QY 1641 AGCACAGCCTTGGGGATATGTTGGAGCGAT 1672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4125 TGCTTGGCCTCGAATCCTTTCCGGTGAGCT 4156

RESULT 12
US-09-960-352-2653
; Sequence 2653, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; NUMBER OF SEQ ID NOS: 2001-09-24
; SEQ ID NO 2653
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 12-LIB3057-020-Q1-K1-C11
US-09-960-352-2653

Query Match      5.7%; Score 102.8; DB 10; Length 415;
Best Local Similarity 56.1%; Pred. No. 2.3e-15;
Matches 194; Conservative 0; Mismatches 152; Indels 0; Gaps 0;

QY 1246 GTCGTGTCATGCTGTGCTCTGCTCTATACCGGAGAAAAATGGTTTCATGGAGCCTGCGG 1305
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 GTCGAACCAATCAGATCCACGTCAGATTCCTGAACAGCTTTTCTACACAAAGCCATTAC 121

QY 1306 TTATTTGTCCTGGGGAATTTTACCCTTTTGGTTCAATCTTTTATTTGAAATGATTATCA 1365
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 CTGGTATTATCATGGGAGGGAATTTGCCCTTTGGTGGCATCTTTATACAGCTTTTCTTCA 181

QY 1366 TCCTCAGCTCTTTCTGGGATATAAGATCTATTATCTCTATGCTTCATGCTGCTGCTGC 1425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 TCCTGAATAGTATTGCTCACACCAAGATGATTACATGTTTGGCTTCTCTCTGCTGGTGT 241
```

```
QY 1426 TGGTTATCTCTGTCATTCGACTGTCTGTGTGACTATTGTGTCACATATTTTCTACTAA 1485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 TTATCAATTTTGGTTATTACCTGTTCTGAAGCAACTATACTCTTTGCTATTTCACCTGT 301

QY 1486 ATGCAGAAATTTACCGGTGGCAATGGACAAGTTTCTCTCTGCTGTCATCAACTGCAATCT 1545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 GTCAGAGGATTTATCATTTGGCAGTGGCGCTCATCTCACCAGCGCTTTACTGCGGTGT 361

QY 1546 ATGTTTACATGTTATCTCTTTACTACTATTATTTTCAAAACAAAGAT 1591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 ATTTCTTAATCTAAGCAATACACTACTCTTCTCAAAACCTGCAAT 407

RESULT 13
US-10-046-935-1582
; Sequence 1582, Application US/10046935
; Patent No. US20020156011A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Secrist, Heather
; APPLICANT: Wang, Aijun
; APPLICANT: Stolk, John A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121-527C1
; CURRENT APPLICATION NUMBER: US/10/046,935
; CURRENT FILING DATE: 2002-01-15
; NUMBER OF SEQ ID NOS: 2239
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1582
; LENGTH: 350
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-046-935-1582

Query Match      5.6%; Score 100.8; DB 9; Length 350;
Best Local Similarity 57.2%; Pred. No. 6.5e-15;
Matches 183; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 1242 CCTTGTCTGTCATCTGCTGCTCTATACCGGAGAAAAATGGTTTCATGGAGCCT 1301
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 CCAGTTTCGAACCAATCAGATTCCACGTCAGATTCTCTGACAGTCGTTCTACACGAAGCCC 87

QY 1302 GCGGTTATTGTTGCTGCTGGTGAATTTTACCTTTTGGTTCAATCTTTTGAATGATAT 1361
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 TTGCTGTTGTTATCATGTTGGAGGGATTTTGGCTTTGGCTGTCATCTTTATACAACTTTTC 147

QY 1362 TTTCATCTTCACGCTCTTCTGGGCATATAAGATCTATTATGCTCTATGCTTCATGCTG 1421
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 TTTCATCTGANTAGTATTTGGTCACACAGATGATTTACATGTTTGGCTTCTCTATTTCTG 207

QY 1422 GTGCTGGTTATCTCTGTCATGTCGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1481
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 GTGTTTATCATTTTGGTTATTTACCTCTCTGAAGCAACATATATCTTTTGGTATTTCCAC 267

QY 1482 CTAAATGCCAGACATTTACCGGTGGCAATGGACAAGTTTCTCTCTGCTGCATCACTGCA 1541
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 CTATGTCGAGAGGATTTATCATTTGGCAATGGCGTTCACTCTTACGAGTGGCTTTACTGCA 327

QY 1542 ATCTATGTTTACATGTTATTC 1561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 GTTTATTTCTTAATCTATGC 347

RESULT 14
US-09-878-178-1582
; Sequence 1582, Application US/09878178
; Patent No. US20020177552A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Harlocker, Susan L.
```

```

: APPLICANT: Secrist, Heather
:
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
:
: TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
:
: FILE REFERENCE: 210121.527
:
: CURRENT APPLICATION NUMBER: US/09/878,178
:
: CURRENT FILING DATE: 2001-06-08
:
: NUMBER OF SEQ ID NOS: 2237
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 1582
:
: LENGTH: 350
:
: TYPE: DNA
:
: ORGANISM: Homo sapien
:
: US-09-878-178-1582

```

Query Match	5.6%	Score 100.8;	DB 9;	Length 350;
Best Local Similarity	57.2%;	Pred. No. 6.5e-15;		
Matches 183;	Conservative 0;	Mismatches 137;	Indels 0;	Gaps 0;
Qy 1242	CCTTGTGCTGTCAATGCTGTGCCCTCATACCGGGAGAAAATAATGTTTCATGGAGCCT	1301		
Db				
Qy 28	CCAGTTCGAACCAATCACAGTTCACAGTCAGATCCTGAAACAGTCGTTCTACAGAAGCCC	87		
Db				
Qy 1302	GCGGTTATTGTTTTGCCGTGGGGAATTTTACCCTTTTGTTCCAATCTTTATTGAAATGTAT	1361		
Db				
Qy 88	TTCGCTGGTATTATCATCGGGAGGGATTTTGGCCTTTGGCTGCATCTTTATACAACCTTTTC	147		
Db				
Qy 1362	TTCACTTCACGCTCTTCTTGGGCATATAAGATCTATTATGTCTATGGCTTCATGATGCTG	1421		
Db				
Qy 148	TTCACTTCGAATAGTATTTTGGTCACACAGATGATTACATGTTTGGCTTCTATTCTG	207		
Db				
Qy 1422	GTGCTGGTTATCCGTGCAATGTGACTGTCTGTGAGCATATTGTGTGCACATATTTTCTA	1481		
Db				
Qy 208	GTGTTTATCATTTTGGTTATTACCTGTCTCTGAAACAATCATACTTCTTGTGTAATTTCCAC	267		
Db				
Qy 1482	CTAAATGCAGAAGATTACCGGTGGCAATGGACAAGTTTCTCTCTGTCTGCAATCAAATGCA	1541		
Db				
Qy 268	CTATGTGCAGAGATTATCATTTGGCAATGGCGTTCATTCCTTACGAGTGGCTTTACTGCA	327		
Db				
Qy 1542	ATCTATGTTTACATGTATTC	1561		
Db				
Qy 328	GTTTATTTCTTAATCTATGC	347		
Db				

```

RESULT 15
US-09-910-943-687
; Sequence 687, Application US/09910943
; Patent No. US2002008160A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Brivanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 687
; LENGTH: 744
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(744)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-687

```

	Query Match	5.3%	Score 99.4;	DB 10;	Length 744;
-	Best Local Similarity	52.6%;	Pred. No. 1.9e-14;		
	Matches 271; Conservative	0;	Mismatches 228;	Indels 16;	Gaps 2;
OY	592	AATATCTGTGTCGCCTCTTTTTCACACATCGGAATCATTTGGTTTTCAATTTCAACTCCT	651		
Dd	100	ATTATNTTTTAGAATCCATGCCTCACACACATTCAGTGTTTAGTATTTATGAATTCGT	159		

Qy	652	TCATGATGGTATCTCTCTGGTGGGCTTGTAGTTTCAATGATATTTAATGAGACATTTAAGAA	711
Db	160	TGSGTATGTGCTCTCCCTTCGGGTATGGTTGCTATGATTTAGTCAAGACATTTACATA	219
Qy	712	AAGATTATGCTCGGTACAGTAAAGAGGAAGAATAATGGATGATATGGATAGAGACCTAGGAG	771
Db	220	ANGATATTGCAAGGTACAATCAGATGGATTCACGAAGATGCTCA-----AG	267
Qy	772	ATGAATATGGATGGAACAGGTGCATGGAGATGCTATTTAGACCATCAAGTCACCCACTGA	831
Db	268	AAGAAATTTGGGTGGAAGCTGGTTCAATGGTGATATTTTCAGAGCACCAAGAAAGGATGC	327
Qy	832	TATTTTCCTCTCTGATTTGGTCTGGATGTCAGATATTTGCTCTGCTCTCATCTGTTATTA	891
Db	328	TGCTCTCTGTTTTCTCGGGTCTTGGGGCTCAGATGCTAATAATGACTTTTGTGCACATTA	387
Qy	892	TTGTTGCAATGATAGAAATTTATATACTGAGAGGGGATCAATGCTCAGTACAGCCATAT	951
Db	388	TTTTTGGCTCCCTTGGATTTTTTCTCCCTGCTAACAGAGGTGCTCTAATGACATGTGCTG	447
Qy	952	TTGTCATATGCTGCTACGTCT---CCAGTGAATGCTTATTTTGGAGGAAGTCTGTATGCT	1007
Db	448	TCCGTGCTGTGGGTGTGCTTGGAACTCCAGCTGGTATGTTGTCTCAAGATTTTACAAA	507
Qy	1008	AGACAAGGAGGAGAGATGGATAAAGCAGATGTTTATTTGGGGCATTCCTTTATCCAGCT	1067
Db	508	TCATTCGGTGGACAAAAGTGGAAACGAATGCTTACTGACTGCATCTCTCTGCCGAGG	567
Qy	1068	ATGGTGTGGGCACTGCCCTCTTTCATCAATTCAT	1102
Db	568	ATTGNAATTCGCTGATTTCTCTTAATGAACATTAAT	602

Search completed: January 20, 2003, 15:45:17  
Job time : 152 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2003, 15:42:01 ; Search time 86 Seconds  
(without alignments)  
6418.816 Million cell updates/sec

Title: US-09-319-724A-13  
Perfect score: 1800  
Sequence: 1 ccgcgcgcgtgtgctgtcgtg.....gtggaaacttgacagcaaaa 1800

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA: \*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq: \*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq: \*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq: \*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq: \*  
5: /cgn2.6/ptodata/1/ina/PCTUS\_COMB.seq: \*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	227.6	12.6	2805	4 US-08-959-004-6	Sequence 6, Appli
c 2	91.6	5.1	769	4 US-09-385-982-530	Sequence 530, Appl
3	51.8	2.9	7218	1 US-08-232-463-14	Sequence 14, Appl
c 4	50.4	2.8	7218	1 US-08-232-463-14	Sequence 14, Appl
5	45.6	2.5	519	1 US-08-686-878A-20	Sequence 20, Appl
6	45.6	2.5	519	4 US-09-175-928-20	Sequence 20, Appl
c 7	44.6	2.5	99500	4 US-09-798-096-10	Sequence 10, Appl
8	42.2	2.3	882	2 US-08-909-965C-9	Sequence 9, Appli
9	38.8	2.2	1707	4 US-09-134-001C-931	Sequence 931, Appl
c 10	38.8	2.2	26664	4 US-09-564-805-28	Sequence 28, Appl
11	38.6	2.1	665	2 US-08-883-795A-36	Sequence 36, Appl
12	38.6	2.1	1001	4 US-09-641-638-103	Sequence 103, Appl
13	38.6	2.1	1847	3 US-08-930-894-3	Sequence 3, Appli
14	38	2.1	4386	4 US-09-300-008B-1	Sequence 1, Appli
c 15	37.8	2.1	8133	1 US-08-480-604A-5	Sequence 5, Appli
c 16	37.8	2.1	8133	2 US-08-405-496A-5	Sequence 5, Appli
=c 17	37.8	2.1	8133	4 US-08-915-136-5	Sequence 5, Appli
c 18	37.8	2.1	8133	4 US-08-957-310-5	Sequence 5, Appli
19	37.6	2.1	447	4 US-09-134-001C-2073	Sequence 2073, Ap
c 20	37.6	2.1	6768	1 US-08-107-755A-1	Sequence 1, Appli
c 21	37.6	2.1	8457	1 US-07-991-867B-1	Sequence 1, Appli
-c 22	37.6	2.1	8457	2 US-08-544-332-1	Sequence 1, Appli
c 23	37.6	2.1	8457	4 US-09-370-861A-1	Sequence 1, Appli
c 24	37.4	2.1	1240	1 US-08-248-466B-6	Sequence 6, Appli
25	37.4	2.1	1510	1 US-08-248-466B-3	Sequence 3, Appli
26	37.4	2.1	1532	1 US-08-248-466B-11	Sequence 11, Appl
27	37.4	2.1	2112	4 US-09-134-001C-178	Sequence 178, Appl

28 36.6 2.0 1609 4 US-09-312-285-1 Sequence 1, Appli  
29 36.6 2.0 1609 4 US-09-728-764-1 Sequence 1, Appli  
c 30 36.4 2.0 959 2 US-08-956-047-35 Sequence 35, Appl  
c 31 36.2 2.0 828 4 US-08-998-416-538 Sequence 538, Appl  
32 36.2 2.0 1383 3 US-09-400-742-7 Sequence 7, Appli  
33 36.2 2.0 1383 3 US-08-618-651A-7 Sequence 7, Appli  
34 36.2 2.0 1383 4 US-09-215-252-6 Sequence 6, Appli  
c 35 36 2.0 19124 2 US-08-487-826B-13 Sequence 13, Appl  
36 35.6 2.0 1094 4 US-09-144-918-3 Sequence 3, Appli  
37 35.6 2.0 1400 4 US-09-144-918-1 Sequence 1, Appli  
38 35.6 2.0 8700 2 US-08-392-625-16 Sequence 16, Appl  
39 35.6 2.0 8700 2 US-08-466-961A-16 Sequence 16, Appl  
40 35.6 2.0 8700 2 US-08-645-193B-18 Sequence 18, Appl  
c 41 35.2 2.0 784 1 US-08-463-115-8 Sequence 8, Appli  
c 42 35.2 2.0 784 1 US-08-465-388-8 Sequence 8, Appli  
43 35.2 2.0 2223 1 US-08-257-073-4 Sequence 4, Appli  
44 35 1.9 2217 4 US-09-134-001C-1448 Sequence 1448, Ap  
45 35 1.9 5433 3 US-08-929-329-1 Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-08-959-004-6  
; Sequence 6, Application US/08959004  
; Patent No. 6197543  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Kaser, Matthew  
; TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE  
; TITLE OF INVENTION: PROTEINS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fastseq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/959,004  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0414 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2805 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: ADREUT06  
; CLONE: 2822412

US-08-959-004-6

```
Query Match 12.6%; Score 227.6; DB 4; Length 2805;
Best Local Similarity 51.9%; Pred. No. 1.5e-50;
Matches 596; Conservative 0; Mismatches 534; Indels 18; Gaps 3;

QY 592 AATATCTTCGATCGCTCTCTTTTCAACATCGGATTCATTTGGTTTTCATTTCAACTCTCT 651
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1044 ACTATATCTGGAGTCTATGCCTCATACCCACATTCAGTGGTTAGCATATATCAATTCCTC 1103

QY 652 TCATGATGGTGATCTCTCTGGGGCTTGTTCATGATGATTTTAAATGAGAAATTAAGAA 711
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1104 TGGTCATTTGCTCTCTCTTATCTGAATGGTAGCTATATGTTTACGGAGACTGCACA 1163

QY 712 AGATTATGCTCGGTACGATAAGAGAGAAATGGATCATATGGATACAGACCTAGGAG 771
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1164 AGATATGCTAGATATATCAGATGGACTCTACGGAAGATGCCAG-----G 1211

QY 772 ATGAATATGGATGGAAACAGGTGCATGGAGATGATTTTATAGACATCAAGTCAACCACCTGA 831
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1212 AAGAAATTTGGCTGGAACTTGTTCATGGTGATATATCCGTCTCCAAGAAAGGATGC 1271

QY 832 TATTTTCTCTCTGATGTTCTTGGATGTCAGATATTTGCTGTGTCCTCTCTCTCTCTCTATTA 891
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1272 TGCATACGCTCTTTAGGATCCGGACACAGATTTTAAATATGACCTTTGTGACTCTAT 1331

QY 892 TTGTTGCAATGATAGAGATTTATATACAGAGAGGGATCAATGCTCAGTAC---AGCCA 948
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1332 TTTTGGCTTGGCTGGATTTTGTACCTTGCCACCGAGGAGCGGTGATGAGCGTGTGCG 1391

QY 949 TATTTTCTCTATGCTGCTAGCTCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTA 1008
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1392 TGGTCTGTGGGTGCTGCTGGGACCCCTGCAGGTATGTTGCTGCCAGATTTCTATAAGT 1451

QY 1009 GACAGGAGAGAGAGATGGATAAGACAGATGTTTATTTGGGGCATTCCTTTATCCAGCTA 1068
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1452 CCTTTGGAGGTGAGAAGTGGAAACAAATGTTTTATTAACATCATTTCTTTGCTCGTGGGA 1511

QY 1069 TGGTGTGGGCACTGCCCTTCTTCATCAATTTTCATAGCCATTTTATACCATGCTTCAAGAG 1128
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1512 TTGATTTTCTGCTGACTTCTTTATTAATGAATCTGATCCTCTGGGGAAGGATCTTCAGCAG 1571

QY 1129 CCATTCCTTTTGGAAACAATTTGGGCGGCTTGTGTGATCTGTTTTTTTGTATTTCTTCCTC 1188
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1572 CTATTCCTTTTGGGACACGGTTGGCATATTTGGCCCTTTTGGTTCTGCATATCTGTGCCCTC 1631

QY 1189 TAAATCTTGTGGTACATCTTGGCCGAAATCTGTACGCTCAGCCCACTTTCCTTGTCTC 1248
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1632 TGACGTTTATTTGGTGATCTCTTTTGGTTTAAAGAAATGCCCATTTGAACAC---CCAGTTC 1688

QY 1249 GTGTCAATGCTGTGCTCTCTATACGGGAGAAATGGTTTCATGGAGCCTGCGGTTA 1308
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1689 GAACCAATCAGATTTCCAGCTCAGATTTCCGACAGTCGTTCTACAGAAAGCCCTTGCCCTG 1748

QY 1309 TTGTTGCTGGGTGGAAATTTTACCTTTTGGTTTCAATCTTTTATTTGAAATGATTTTCATCT 1368
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1749 GTATTATCATGGAGGATTTTGGCCCTTTGGCTGATCTTTATACAACTTTTCTTCATTC 1808

QY 1369 TCAGCTCTTTCTGGGCATATATGATCTATTTATGCTATGCTGCTTATGATGCTGGTGGTGG 1428
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1809 TGAATAGTATTTTGGTCACACCAAGATGATTTACATGTTTGGCTTCTCTCTCTCTCTCTCT 1868

QY 1429 TTATCTCTGTGCAATCTGACTGTCTGTGACATTTTGTGACATATTTTCTCTCTCTCTCT 1488
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1869 TCATTTTGGTTATTTACCTGTTCTGAAGCAACTATCTCTTTTGGTATTTTCCACCTATGTTG 1928

QY 1489 CAGAAGATTACCGGTGGCAATGGCAAGTTTCTCTCTGCTGCATCAACTGCAATCTATG 1548
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1929 CAGAGGATTTATTTGGCAATGGCTTCTATTTCTTACGAGTGGCTTTACTGAGTTTAT 1988

QY 1549 TTTTACATGATTTCTTTTACTACTATTTTTCATAAACAAGATGATGCGTTATTTTCAAA 1608
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 1989 TCTTAATCTATGCATACACTACTTCTTTTCAAAACTGCAGATCATCGGGAACAGCA 2048
```

```
QY 1609 CATCATTTTACTTTTGGATATATGCGGTATTTAGCACAGCCTTTGGGATAATGTGTGGAG 1668
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 2049 CAATTCTGTACTTTGGTTATACCATGATAATGGTTTGTGATCTTCTTTTACAGGAA 2108

QY 1669 CGATTGGTTTACATGGGAACAAGTGGCTTTGTCCGAAAAATCTATCTAATGTGAAAAATTG 1728
D 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111 1111
Db 2109 CAATTGGCTTCTTTCATGCTTTTGGTTTGTGTTTACCAAAATATACAGTGTGGTGAAGGTTG 2168

QY 1729 ACTAGAGA 1736
D 111 111
Db 2169 ACTGAAGA 2176
```

RESULT 2

```
US-09-385-982-530/c
; Sequence 530, Application US/09385982
; Patent No. 5262334
; GENERAL INFORMATION:
; APPLICANT: ENDEGE, WILSON O., ET AL.
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE OF INVENTION: PRODUCTS: 11
; FILE REFERENCE: CCDA-260XX
; CURRENT APPLICATION NUMBER: US/09/385,982
; EARLIER FILING DATE: 1999-08-30
; EARLIER APPLICATION NUMBER: 09/328,111
; EARLIER FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: 60/117,393
; EARLIER FILING DATE: 1999-01-27
; EARLIER APPLICATION NUMBER: 60/098,639
; EARLIER FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 544
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 530
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(769)
; OTHER INFORMATION: n = A,T,C or G
US-09-385-982-530
```

```
Query Match 5.1%; Score 91.6; DB 4; Length 769;
Best Local Similarity 57.0%; Pred. No. 7.1e-15;
Matches 166; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
```

```
QY 1261 TGCCTCGTCTATACCGGAGAAAAATGTTTCATGGAGCCTGCGGTATTTGTTGCCCTGG 1320
D 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 308 TCCACNGTCAGATTCCTGAACAGTGTCTACACGAAGCCCTTGCCTGATTTATCATGG 249

QY 1321 GTGGAATTTTACCTTTTGGTTCAATCTTTATTTGAATGATTTTCATCTTCAGCTCTTTCT 1380
D 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 248 GAGGATTTTGGCCCTTTTGGCTGCATCTTTTCAACAACTTTTCTTCATCTGAAATGATTT 189

QY 1381 GGGCATAAAGATCTATTATGCTTATGCTTTCATGCTGCTGCTGCTGTTATCTCTGCA 1440
D 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 188 GTACACACAGATGTTTACATGTTTGGCTTCTCTATTTCTGTTGTTATCATTTTGGTTA 129

QY 1441 TTGTGACTGTCTGTGACTATTTGTGTGCACATATTTTCTACTAAATGCAGAAATACC 1500
D 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 128 TTACCTGTTCTGAAGCAACTATATCTTCTTGTCTATTTCCACCTATGTGCAGAGGATTATC 69

QY 1501 GGTGGCAATGGACAAGTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1551
D 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 68 ATTGGCAATGGCGTTTCATTCCTTTACGAGTGGCTTTTACTGCACTTTATTCTT 18
```

RESULT 3

```
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
```

APPLICANT: DORNER, F.  
 APPLICANT: SCHEIFLINGER, F.  
 APPLICANT: FALKNER, F. G.  
 TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
 NUMBER OF SEQUENCES: 52  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 1800 Diagonal Road, Suite 500  
 CITY: Alexandria  
 STATE: VA  
 COUNTRY: USA  
 ZIP: 22313-0299  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/232,463  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/07/935,313.  
 FILING DATE:  
 APPLICATION NUMBER: EP 91 114 300.6  
 FILING DATE: 26-AUG-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)836-9300  
 TELEFAX: (703)683-4109  
 TELEX: 899149  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 7218 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 CLONE: pTZgpt-Fls  
 PS-08-232-463-14

Query Match	2.9%;	Score 51.8;	DB 1;	Length 7218;	
Best Local Similarity	4.7%;	Pred. NO. 0.0006;			
Matches	17;	Conservative	202;	Mismatches 144;	Indels 0; Gaps 0;

  

QY	1053	TTCCTTATCCAGCATTGGTGTTGGCACGTCCTCTTCATCAATTCATAGCCCATTTAT	1112
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
Db	1104	YY	1163
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
QY	1113	TACCATCGTTCAGAGGCATTCCTTTTGGAAACAATGGTGCCGTGTGTGCATCTGTTTT	1172
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
Db	1164	YY	1223
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
QY	1173	TTTGTATTCTTCCTCTAATACTTGTGGTACAACTACTGGCGGAATCTCTCAGGTCA	1232
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
Db	1224	YY	1283
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
QY	1233	CCCAACTTCTTGTCGTGTCAATGCTGTGCCCTCTATACGGGAGAAAAATGGTTC	1292
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
<b>Db</b>	1284	YY	1343
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
QY	1293	ATGSAGCCTGGGTATTGTGTGCTGGGTGGAATTTACCTTTTGGTTCATCTTATT	1352
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
<b>Db</b>	1344	YY	1403
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
QY	1353	GAAATGTAATTCATCTCACGTCCTTCTGGGCATATAAGATCTATTATGCTATCGGCTTC	1412
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
Db	1404	YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGTACCAAATTCCTCTATCTCTTAACTAC	1463
	:	: : : :	:
	:	: : : :	:
	:	: : : :	:
QY	1413	ATG 1415	

Db	1464	TTG	1466
----	------	-----	------

RESULT 4  
US-08-232-463-14/c  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; APPLICANT: FALKNER, F. G.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEFAX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: pTZ19pt-F15  
; US-08-232-463-14

Query Match	2.8%;	Score 50.4;	DB 1;	Length 7218;
Best Local Similarity	4.5%;	pred. NO. 0.0014;		
Matches	18;	Conservative 219;	Mismatches 165;	Indels 0; Gaps 0;
QY	166	GGTCAAAAAAGATACAGTCATTACCATGAACCTCTGGGAGAACCACTTCAACGGGTTG	225	
Db	1460	GTTAAACAGATAGAAGAAATTTCGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRR	1401	
QY	226	AATTGGAAATTTAGTGGTCTGGATATTAAATTTAAAGATGATGTGATGCCACCACTTACT	285	
Db	1400	RR	1341	
QY	286	GTGAATTTGATTTAGATTAAGAAAAAGAGAGATGCATTTGTATATGCCATAAAAATCATTT	345	
Db	1340	RR	1261	
QY	346	ACTGGTACCAGATGATACATGATTTACCAATATGGGGTATTGTTGGTGAGCGCTGATG	405	
Db	1280	RR	1221	





FEATURE;  
NAME/KEY: unsure  
LOCATION: (157)  
FEATURE;  
NAME/KEY: unsure  
LOCATION: (161)  
FEATURE;  
NAME/KEY: unsure  
LOCATION: (204)  
FEATURE;  
NAME/KEY: unsure  
LOCATION: (239)  
FEATURE;  
NAME/KEY: unsure  
LOCATION: (305)  
US-09-175-928-20

Query Match 2.5%; Score 45.6; DB 4; Length 519;

Best Local Similarity 47.5%; Pred. No. 0.0085;  
Matches 160; Conservative 0; Mismatches 174; Indels 3; Gaps 1;

QY 769 GAGATGAATATGATGGAACAGTGCATGAGATGATTTAGACCATCAAGTCAACCCAC 828  
DB 149 GGAAGAAATNCTGGAACCTGTTCAATGGTGATATATACCTCTCCCAAGAAANGGA 208  
QY 829 TGATATTTCTCTCTGATGTTCTGGATGTCAGATATTTGCTGTCCTCATCGTTA 888  
DB 209 TCGTCTATCAGTCTTTCTAGGAGCCGGGANACAGATATTAATTATGACCTTTGTGACTC 268  
QY 889 TTATTGTCGAATAGTAGAATTTATATCTAGAGGGGATCAATGCTCAGTACA---G 945  
DB 269 TATTTTCGCTCGCTGGGAGTTTGTACCTCCCAAGGAGCGCTGATGACGTGTG 328  
QY 946 CCATATTTGCTATGCTGCTACCTCTCCAGTGAATGGTATTTTGGAGGAAGTCTGTATG 1005  
DB 329 CTGTGCTCTGTGGTGTCTGGGACCCCTCGAGGCTATGTTTCGCAGATTTCTATA 388  
QY 1006 CTAGACAAGGAGGAGATGATGAAGCAGATGTTTATTTGGGGATTCCTTATCCCGAG 1065  
DB 389 AGTCCTTTGGAGTGAGAGTGGAACAAATGTTTTATTAAACATCATTTCTTTGTCCTG 448  
QY 1066 CTATGCTGTGCGCACTGCTCTTCTTCATCAATTCAT 1102  
DB 449 GGATTGATTTGCTGACTTCTTTTATAATGAATCTGAT 485

## RESULT 7

US-09-798-096-10/c  
Sequence 10, Application US/09798096  
Patent No. 6399378

GENERAL INFORMATION:  
APPLICANT: Donna T. Ward  
TITLE OF INVENTION: ANTISENSE MODULATION OF RECL2 EXPRESSION  
FILE REFERENCE: R15-0207  
CURRENT APPLICATION NUMBER: US/09/798,096  
CURRENT FILING DATE: 2001-03-01  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 10

LENGTH: 99500  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:

US-09-798-096-10

Query Match 2.5%; Score 44.6; DB 4; Length 99500;

Best Local Similarity 46.3%; Pred. No. 0.15;  
Matches 146; Conservative 0; Mismatches 169; Indels 0; Gaps 0;

QY 507 AAACGTGTTCCAAATCTAAATCCAGATGTCATATTCAGTAAATGGAAAGTCAGAT 566  
DB 9936 ATACAGTTTCAATAGCTAGAGGAGGATATTGAATGTCCTCCCAACACAAAGAAATGATAA 9877

QY 567 GTGAAATTTGAAGATGATTTGACAAATATCTTGATCCGTCCTTTTTCACATCGGATT 626  
DB 9876 ATGTTTTAAGATGATAGATATGCTAATTAACCTGATCATATACAGTATATGTAT 9817  
QY 627 CATTTGGTTTTCAATTTTCAACTCCTTCATGATGCTCTTGGTGGGCTTAGTTTCA 686  
DB 9816 CAAAATATCACTATGTACCCCAATAAATTTGTTGGGTACAAATTTATGTCACCTTTAAGAAAT 9757  
QY 687 ATGATTTTAATGAGAACATTTAAGAAAAAGATTATGCTCGGTACAGTAAAGAGGAAGAAATG 746  
DB 9756 GTAAAAATAAAAAAGTCATTTGACTAAAAATTATGTTAAATFCCACTTAAAAAGCCAAAGTA 9697  
QY 747 GATGATATGATAGAGACCTAGGAGATGAATATGATGGAACAGGTGCGATGCGAGATGTA 806  
DB 9696 GAAGATATCTGTACACCTATGTAATGTCTATGATGATATATATAAATGTAGATAAC 9637  
QY 807 TTTAGACCATCAAGT 821  
DB 9636 TATATATAAATATGT 9622

## RESULT 8

US-08-909-965C-9  
Sequence 9, Application US/08909965C  
Patent No. 5936078

GENERAL INFORMATION:  
APPLICANT: Kuga Tetsuo  
APPLICANT: Nakagawa Satoshi  
APPLICANT: Sakaki Yoshiyuki  
APPLICANT: Zhao Nanding  
APPLICANT: Hashida Hideji  
TITLE OF INVENTION: NOVEL DNA, NOVEL POLYPEPTIDE  
TITLE OF INVENTION: AND NOVEL ANTIBODY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO  
STREET: 277 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10172-0194

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/909,965C  
FILING DATE: August 12, 1997  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 322745/95  
APPLICATION NUMBER: PCT/JP96/03630  
FILING DATE: 12-NO. 5936078-1995  
FILING DATE: 12-Dec-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lawrence S. Perry  
REGISTRATION NUMBER: 31865  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-2400  
TELEFAX: 212-758-2982  
TELEX: 236262

INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 882 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE: human  
ORGANISM: human  
IMMEDIATE SOURCE:  
CLONE: F1180

```

; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97 to 195
; IDENTIFICATION METHOD: by experiment
; US-08-909-965C-9
; Query Match
; Best Local Similarity 2.3%; Score 42.2; DB 2; Length 882;
; Matches 95; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1554 ATGTATTCCTTTTACTACTATTTTTCACAAAGAGATGATGGCTTATTTTCAACACATCA 1613
Db 19 ATCTATGCAGTACACTACTCTCTTTTCAAACTGACAGTACACGGGAAACAGACACAAT 78

QY 1614 TTTTACTTTGGATATATGCGGCTATTTAGCACAGCCTTGGGGATATGTTGGAGCGATT 1673
Db 79 CTGTACTTTGGTATACCATGATAATGTTTGTGATCTCTTCTTTTACAGGAACAAT 138

QY 1674 GGTATACATGGGAACAGTGCCTTTTCCGAAAAATCTATACTAATGTGAAAAATTTGACTAG 1733
Db 139 GCCTCTTTTGCATGCTTTTGGTTTCTTACCAAAATATACAGTGTGTGAAGTTGACTCA 198

QY 1734 AGA 1736
Db 199 AGA 201

RESULT 9
US-09-134-001C-931
; Sequence 931, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 931
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-931

Query Match 2.2%; Score 38.8; DB 4; Length 1707;
Best Local Similarity 43.5%; Pred. No. 0.89;
Matches 175; Conservative 0; Mismatches 227; Indels 0; Gaps 0;

QY 202 TGGGAGAGCCTTCAAGGGTTGAATTTAGTGGCTGGATATTAATTTAAAG 261
Db 500 TTGGTGAAGAGCTGAACCACTTGAATCGAAATCGAAATTTATGAGCCCAAGTTAAATG 559

QY 262 ATGATGTGATGCCACCCACTTACTCTGAAATTTAGATTAAGAAAGAGAGATGCAT 321
Db 560 AATATGAGAAATTTAAAGTGAAGTAAATATGTTCAAGCTCATTAATCATGTCGCT 619

QY 322 TTGTATGCCATAAAATCAATTTACTGTCACAGATGATACATGATTTTACCAATAT 381
Db 620 TAGAAGATCAAAATTTAAATTTAAATCTTATATGATGATCAATTTCCAGAAATTAATTCGAG 679

QY 382 GGGGTATTTGGTGAGGCTGATGAAATGGAGAGATTTACTATCTTTGGACCTATAAAA 441
Db 680 AAACCTCAAAAAGAAATTTACCAGGACAGTTTCAAGATTTTAAATATGATGATGATGAGACTTAA 739

QY 442 AACTTGAATAGTTTAAATGAAATCGAATTTGTTGATGTTAACTTAATCTACTAGTGAAGAA 501
Db 740 AGGTTGAAGGCTATGATTTAGACCATGTCGAAGTGCATGATGATGATGATGATGATGATGAT 799

```

```

QY 502 AGGTGAAACTGGTTCCTCAAAATCTAAATCCAGATGTCATATTCAGTAAATGAAAAAGT 561
Db 800 CAGAAATTAAGTTTGTGTAACCAATGATTAGTCGATTAGAAATAGATGAAGCTAATAATA 859

QY 562 CAGATGTCGAAATTTGAAGATCGATTTTGACAAATATCTTTGATC 603
Db 860 AACTTTGAAATATTAATGATAAAATTAGATAAATGATGATGATGATGATGATGATGATGAT 901

RESULT 10
US-09-564-805-28/c
; Sequence 28, Application US/09564805
; Patent No. 6333403
; GENERAL INFORMATION:
; APPLICANT: Tavtigian, Sean V.
; APPLICANT: Teng, David H.F.
; APPLICANT: Simard, Jacques
; APPLICANT: Rommens, Johanna M.
; APPLICANT: Myriad Genetics, Inc.
; TITLE OF INVENTION: Chromosome 17p-Linked Prostate Cancer Susceptibility
; FILE REFERENCE: 2318-258
; CURRENT APPLICATION NUMBER: US/09/564,805
; CURRENT FILING DATE: 2000-05-05
; PRIOR APPLICATION NUMBER: US 60/107,468
; PRIOR FILING DATE: 1998-11-06
; PRIOR APPLICATION NUMBER: 09/434,382
; PRIOR FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 240
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 26664
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (910)..(13104)
; OTHER INFORMATION: exon 1: 910-1154; exon 2: 1736-1786; exon 3:
; OTHER INFORMATION: 1925-1995; exon 4: 3025-3089; exon 5: 4361-4418;
; OTHER INFORMATION: exon 6: 5582-5650; exon 7: 7075-7194; exon 8:
; OTHER INFORMATION: 8186-8244; exon 9: 12878-12936; exon 10:
; NAME/KEY: misc_feature
; LOCATION: (13756)..(22917)
; OTHER INFORMATION: exon 11: 13756-13868; exon 12: 15283-15378; exon
; OTHER INFORMATION: 13: 16278-16416; exon 14: 16498-16583; exon 15:
; OTHER INFORMATION: 18583-18701; exon 16: 20349-20445; exon 17:
; OTHER INFORMATION: 22172-22310; exon 18: 22879-22917
; NAME/KEY: misc_feature
; LOCATION: (23045)..(26452)
; OTHER INFORMATION: exon 19: 23045-23154; exon 20: 23795-23895; exon
; OTHER INFORMATION: 21: 23973-24093; exon 22: 24354-24432; exon 23:
; OTHER INFORMATION: 25026-25170; exon 24: 25812-26036; polyadenylation
; OTHER INFORMATION: signal: 26447-26452
; NAME/KEY: variation
; LOCATION: (826)..(23879)
; OTHER INFORMATION: s at positions 826 and 23180 is G or C; y at
; OTHER INFORMATION: positions 1914, 5568, 7165, 16431, 1857 and 20486
; OTHER INFORMATION: is C or T; n at position 13128 is t or tgat; r at
; OTHER INFORMATION: positions 22211 and 23879 is A or G.
US-09-564-805-28

Query Match 2.2%; Score 38.8; DB 4; Length 26664;
Best Local Similarity 48.9%; Pred. No. 2.9;
Matches 133; Conservative 0; Mismatches 137; Indels 2; Gaps 1;

QY 436 ATAAAAACTTGAATAGGTTTTTAATGGAATCGAATTTGTTGATGTTTAACTTAATGATG 495
Db 11736 AAGAAAAAAGAAAAAACCCTTAATATGCTATTACATTTAGAAAGGCAACAAATATTTTAC 11677

QY 496 AAGGAAAGGTGAACTGGTTCCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCA 555
Db 11676 CAGAAACTTCACATCAATTTTATAGGAAAAAATAATTTTCCACAAACAAAAAATGTC 11617

```





;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/329,154  
;; FILING DATE: 25-OCT-1994  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/161,907  
;; FILING DATE: 02-DEC-1993  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/985,321  
;; FILING DATE: 04-DEC-1992  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/429,791  
;; FILING DATE: 31-OCT-1989  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: INGOLIA, DIANE E.  
;; REGISTRATION NUMBER: 40,027  
;; REFERENCE/DOCKET NUMBER: OPHD-01763  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (415) 705-8410  
;; TELEFAX: (415) 397-8338  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 8133 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 1..8130  
US-08-480-604A-5

Query Match 2.1%; Score 37.8; DB 1; Length 8133;  
Best Local Similarity 44.2%; Pred. No. 3.2;  
Matches 156; Conservative 0; Mismatches 197; Indels 0; Gaps 0;  
QY 355 AGATGTACATAGATGATTACCAATATGGGGTATTGTTGGTGAGGCTGATGAAAAATCGAG 414  
Db 4210 ATATATATCTATCTTTTATTATCTATATATCGCCTGAAAAATCTATTGTTGATTCCTTATA 4151  
QY 415 AAGATTACTATCTTTGGACCTATAAAAAAGTGAATAGTGTATTAATGGAATCGAATTG 474  
Db 4150 TAGTTTATTTTATTATATCAATTTTACTTAAACATCTTTTATTAACTTTTCCTTTT 4091  
QY 475 TTGATGTTAATCTACTAGTAGAGAAAGTGAAAGTGGTTCCAAATCTAAAAATCCAGA 534  
Db 4090 TAATAGTACCATTTTCTATAGATATTTCTTACTTCAATATCAATATTAATATCCATA 4031  
QY 535 TGTCATATTTCAGTAAATGGAAGTCAAGATGTAATTTGAAGATCGATTTGACAAAT 594  
Db 4030 AATCATCTTTAGATAAATTTATATTCGTTGATATGGATGAAGATAATAATAAAGAGT 3971  
QY 595 ATCTTGATCCGCTCTTTTTCACATCGGATTCATTTGTTTCAATTTTTCACATCCTTCA 654  
Db 3970 AGTTCCCTCCTGCTCCATCAATGAATAAGATAAATTTGTTCTTAATTTCTAGTAGTTA 3911  
QY 655 TGATGATGATCTCTTTGGTGGGCTTAGTTTCAATGATTTTAATGAGAACATTA 707  
Db 3910 TAGTTGGCATTAAGAAGTTTCTAGTAGTATCTTATCTAGTTTAAATTTAATATTA 3858

Search completed: January 20, 2003, 16:35:31  
Job time : 356 secs











QY 1141 GAACAATGGTGGCGTTTGTGGCATCTGTTTGTATTTCTTCCCTAAATCTTGTG 1200  
DB 1205 GAACAATGGTGGCGTTTGTGGCATCTGTTTGTATTTCTTCCCTAAATCTTGTG 1264  
QY 1201 GTACAATACTTGGCCGAAATCTGTACAGTGCAGCCAACTTTCCTGTGCTCAAGCTG 1260  
DB 1265 GTACAATACTTGGCCGAAATCTGTACAGTGCAGCCAACTTTCCTGTGCTCAAGCTG 1324  
QY 1261 TGGCTCTCTCTATACCGAGAGAAAATGGTTCATGGAGCCTGCGGTTATTGTTTGCCTGG 1320  
DB 1325 TGGCTCTCTCTATACCGAGAGAAAATGGTTCATGGAGCCTGCGGTTATTGTTTGCCTGG 1384  
QY 1321 GTGGAATTTTACCTTTTGGTTCATCTTTATTAATAATGATTTTCATCTTCACGCTCTTTCT 1380  
DB 1385 GTGGAATTTTACCTTTTGGTTCATCTTTATTAATAATGATTTTCATCTTCACGCTCTTTCT 1444  
QY 1381 GGGCATATAGATCTATATGTCTATGGCTTCATGATGCTGGTGGTGTATCTGTGCA 1440  
DB 1445 GGGCATATAGATCTATATGTCTATGGCTTCATGATGCTGGTGGTGTATCTGTGCA 1504  
QY 1441 TTGTGACTGCTGTGCTACTATTGTGTCACATATTTTCTACTAAATGCAGAGATTACC 1500  
DB 1505 TTGTGACTGCTGTGCTACTATTGTGTCACATATTTTCTACTAAATGCAGAGATTACC 1564  
QY 1501 GGTGGCAATGGACAAGTTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATGTATT 1560  
DB 1565 GGTGGCAATGGACAAGTTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATGTATT 1624  
QY 1561 CCTTTTACTACTATTTTCAAAACAAGATGATGCTTATTTCAACATCATTTTACT 1620  
DB 1625 CCTTTTACTACTATTTTCAAAACAAGATGATGCTTATTTCAACATCATTTTACT 1684  
QY 1621 TTGGATATATGGCGTATTTAGCACAGCCTTGGGATAATGTTGGAGCATTGGTTACA 1680  
DB 1685 TTGGATATATGGCGTATTTAGCACAGCCTTGGGATAATGTTGGAGCATTGGTTACA 1744  
QY 1681 TGGGAACAAGTGGCTTTGTCCGAAAATCTATCTATATGTAATTTGACTAGAGACCCA 1740  
DB 1745 TGGGAACAAGTGGCTTTGTCCGAAAATCTATCTATATGTAATTTGACTAGAGACCCA 1804  
QY 1741 AGAAACCTGGAATTTGGATCAATTTCTTTTCATAGGGGTGGAATTCGACAGCAAAA 1800  
DB 1805 AGAAACCTGGAATTTGGATCAATTTCTTTTCATAGGGGTGGAATTCGACAGCAAAA 1864

RESULT 3

ID AAA16630 standard; cDNA; 3370 BP.

AC AAA16630:

XX 16-JUN-2000 (first entry)

XX Human secreted protein clone pk366.7 nucleotide sequence SEQ ID NO:25.  
DE Human; secreted protein; immunostimulant; immunosuppressant; virucide;  
KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;  
KW antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide;  
KW antihypoid; immune deficiency; severe combined immunodeficiency; SCID;  
KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;  
KW connective tissue disease; multiple sclerosis; erythematosis;  
KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;  
KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
KW insulin dependent diabetes mellitus; graft-versus-host-disease;  
KW autoimmune inflammatory eye disease; allergy; ss.

XX Homo sapiens.

XX WO200009552-A1.

XX 24-FEB-2000.

XX

PF 13-AUG-1999; 99WO-US18298.  
XX 14-AUG-1998; 98US-0096622.  
PR 17-AUG-1998; 98US-0096815.  
PR 04-SEP-1998; 98US-0099229.  
PR 23-OCT-1998; 98US-0105368.  
PR 08-JAN-1999; 99US-0115234.  
PR 12-FEB-1999; 99US-0119931.  
PR 18-FEB-1999; 99US-0120575.  
PR 30-APR-1999; 99US-0132020.  
PR 11-AUG-1999; 99US-0096622.  
XX (GEMY ) GENETICS INST INC.  
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;  
PI Wong GG, Clark HF, Fechtel K;  
DR WPI; 2000-205979/18.  
DR P-PSDB; AAY94910.  
XX New polynucleotides encoding secreted proteins, which may have e.g.  
PT nutritional, chemokine, immune stimulating or suppressing,  
PT hematopoiesis regulating, tissue growth, activin/inhibin  
PT antiinflammatory or tumor inhibition activity  
XX Claim 34; Page 493; 641pp; English.  
XX AAA16618 to AAA16697 encode the human secreted proteins given in  
CC AAY94898 to AAY94980, isolated from human adult brain, adult thyroid,  
CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,  
CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,  
CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,  
CC and adult bladder, cDNA libraries. The polynucleotides and proteins are  
CC predicted to have biological activities which would make them suitable  
CC for treating, preventing or ameliorating medical conditions in humans  
CC and animals. The polynucleotides can be used as markers for tissues in  
CC which the protein is preferentially expressed, as molecular weight  
CC markers on Southern gels, and as chromosome markers or tags to identify  
CC chromosomes or to map gene positions. The proteins can be used in the  
CC treatment of immune deficiencies and disorders, such as severe combined  
CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other  
CC infections. These infections include human immunodeficiency virus (HIV),  
CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and  
CC candidiasis. The proteins can be used to treat autoimmune disorders such  
CC as connective tissue disease, multiple sclerosis, systemic lupus  
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,  
CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent  
CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and  
CC autoimmune inflammatory eye disease. The proteins can also be used to  
CC treat allergic conditions, such as asthma. AAA16698 to AAA16774 represent  
CC probes for the human secreted proteins from the present invention.  
XX Sequence 3370 BP; 1031 A; 537 C; 676 G; 1126 T; 0 other;  
SQ

Query Match 99.7%; Score 1794.4; DB 21; Length 3370;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1795; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 CGCGCTGTGGCTGCTGCTGCTGCTGCCCGGACCGGGCGGAGCAGCAGAACAC 64  
DB 1 CGCGCTGTGGCTGCTGCTGCTGCTGCCCGGACCGGGCGGAGCAGCAGAACAC 60  
QY 65 GTATCAAGATAAAGAGAGACTTCTCTTATGGATGAATCTGTTGGGCCCTACCATAATCG 124  
DB 61 GTATCAAGATAAAGAGAGAGTGTCTTATGGATGAATCTGTTGGGCCCTACCATAATCG 120  
QY 125 TCAAGAAACATATAAGTACTTTTCACTTCTCTGTGGGTCAAAAAAGATATGAC 184  
DB 121 TCAAGAAACATATAAGTACTTTTCACTTCTCTGTGGGTCAAAAAAGATATGAC 180  
QY 185 TCATTACCATGAACCTCTGGGAGAACACTTCAAGGGTGTGAATTGGAATTAGTGTCT 244  
|||||

Db 181 TCATTACCATGAACACTCTGGGAGAAGACATTCAGGGGTTGAATTGGAAATTTAGTGTCT 240  
QY 245 GGATATTAAATTTAAAGATGATGTGATGCCAGGCACCTTACTGTGAAATTCATTTAGATAA 304  
Db 241 GGATATTAAATTTAAAGATGATGTGATGCCAGGCACCTTACTGTGAAATTCATTTAGATAA 300  
QY 305 AGAAAAGAGAGATGCATTTGTATATGCCATAAAAAATCATTTACTGGFPACAGATGTACAT 364  
Db 301 AGAAAAGAGAGATGCATTTGTATATGCCATAAAAAATCATTTACTGGFPACAGATGTACAT 360  
QY 365 AGATGATTTTACCATAATATGGGGTATTTGTGTGAGGCTGATGAAATCGCAAGATTTACTA 424  
Db 361 AGATGATTTTACCATAATATGGGGTATTTGTGTGAGGCTGATGAAATCGCAAGATTTACTA 420  
QY 425 TCTTTGGACCTATAAAAACTTCAATAGGTTTTTAATGGAATTCGAATTCCTTGATGTTAA 484  
Db 421 TCTTTGGACCTATAAAAACTTCAATAGGTTTTTAATGGAATTCGAATTCCTTGATGTTAA 480  
QY 485 TCTAACTAGTGAAGGAAGGTGAACGTGTTTCCAAATFACTAAATCCAGATGTCAATTC 544  
Db 481 TCTAACTAGTGAAGGAAGGTGAACGTGTTTCCAAATFACTAAATCCAGATGTCAATTC 540  
QY 545 AGTAAATGAAAAGTCAGATGTGAATTTGAAGATCGATTTTGACAAATATCTTGATCC 604  
Db 541 AGTAAATGAAAAGTCAGATGTGAATTTGAAGATCGATTTTGACAAATATCTTGATCC 600  
QY 605 GTCCCTTTTTCACATCGGATTCATTTGGTTTTTCAATTTTCAACTCCCTTCATGATGTGAT 664  
Db 601 GTCCCTTTTTCACATCGGATTCATTTGGTTTTTCAATTTTCAACTCCCTTCATGATGTGAT 660  
QY 665 CTTCTTGGTGGCTTAGTTTCAATGATTTTAAATGAGAACAATTAAAGAAAAGATTATGCTCG 724  
Db 661 CTTCTTGGTGGCTTAGTTTCAATGATTTTAAATGAGAACAATTAAAGAAAAGATTATGCTCG 720  
QY 725 GTACAGTAAAGAGAGAANAATGATATGATAGAGACCTAGGAGATGAATATGGATG 784  
Db 721 GTACAGTAAAGAGAGAANAATGATATGATAGAGACCTAGGAGATGAATATGGATG 780  
QY 785 GAAACAGGTGTCATGGAGATGATTTAGACCATCAAGTCACCCACTGATATTTTCCCTCTCT 844  
Db 781 GAAACAGGTGTCATGGAGATGATTTAGACCATCAAGTCACCCACTGATATTTTCCCTCTCT 840  
QY 845 GATTGGTCTCGATGTCAGATATTTGCTGTCTCTCATCTGTTATTTATGTTTCCAATGAT 904  
Db 841 GATTGGTCTCGATGTCAGATATTTGCTGTCTCTCATCTGTTATTTATGTTTCCAATGAT 900  
QY 905 AGAAGATTTATATCTACAGAGGGATCAATGCTCAGTACAGCCATATTTCTCTATGCTGC 964  
Db 901 AGAAGATTTATATCTACAGAGGGATCAATGCTCAGTACAGCCATATTTCTCTATGCTGC 960  
QY 965 TACGTCTCCAGTGAATGGTTTATTTTGGAGGAAGTCTGTATGCTAGACAAGGAGGAGGAG 1024  
Db 961 TACGTCTCCAGTGAATGGTTTATTTTGGAGGAAGTCTGTATGCTAGACAAGGAGGAGGAG 1020  
QY 1025 ATGGATAAGCAGATGTTTATTTGGGCAATTCCTTATCCAGCTATGTTGTGGCACTGC 1084  
Db 1021 ATGGATAAGCAGATGTTTATTTGGGCAATTCCTTATCCAGCTATGTTGTGGCACTGC 1080  
QY 1085 CTTCTTTCATCAATTTTCATAGCCATTTATTTACCATGCTTCAAGAGCCATTCCTTTTGAAC 1144  
Db 1081 CTTCTTTCATCAATTTTCATAGCCATTTATTTACCATGCTTCAAGAGCCATTCCTTTTGAAC 1140  
QY 1145 AATGGTGGCGGTTGTGTGATCTGTTTTTTTGTATTCTTCCCTCAATCTTCTGTTGATC 1204  
Db 1141 AATGGTGGCGGTTGTGTGATCTGTTTTTTTGTATTCTTCCCTCAATCTTCTGTTGATC 1200  
QY 1205 AATPACTTGGCCGAAATCTGTGAGGTGAGCCCAACTTTTCCCTTGTGCTCAATGCTGTGCC 1264  
Db 1201 AATPACTTGGCCGAAATCTGTGAGGTGAGCCCAACTTTTCCCTTGTGCTCAATGCTGTGCC 1260  
QY 1265 TCGTCTCTATACCGGAGAAAATGGTTTCATGGAGCCTGGGTTATTTGTTGCTGGGTGG 1324  
Db 1261 TCGTCTCTATACCGGAGAAAATGGTTTCATGGAGCCTGGGTTATTTGTTGCTGGGTGG 1320

QY 1325 AATTTTACCCTTTGGTTCAATCTTTATTGAAATCTATTTCATCTTCACTGCTTTCTGGGC 1384  
Db 1321 AATTTTACCCTTTGGTTCAATCTTTATTGAAATCTATTTCATCTTCACTGCTTTCTGGGC 1380  
QY 1385 ATAAAGATCTATTATGCTCTATGCTTTCATGCTGCTGGTGTATCTGCTGCTGTCATTT 1444  
Db 1381 ATAAAGATCTATTATGCTCTATGCTTTCATGCTGCTGGTGTATCTGCTGTCATTTG 1440  
QY 1445 GACTGCTGTGCTGACTATTGTTGTCACATATTTTCTACTAAATGCGAAGATTTACCGGTG 1504  
Db 1441 GACTGCTGTGCTGACTATTGTTGTCACATATTTTCTACTAAATGCGAAGATTTACCGGTG 1500  
QY 1505 GCAATGACAAAGTTTCTCTCTGCTGCTCACTCACTGCAATCTATGTTTACATGATTTCTCT 1564  
Db 1501 GCAATGACAAAGTTTCTCTCTGCTGCTCACTCACTGCAATCTATGTTTACATGATTTCTCT 1560  
QY 1565 TTACTACTATTTTTCAAAAAAGATGATGCTTATTTTCAACATCAATTTTACTTTTGG 1624  
Db 1561 TTACTACTATTTTTCAAAAAAGATGATGCTTATTTTCAACATCAATTTTACTTTTGG 1620  
QY 1625 ATATATGGCGGTATTTAGCACAGCCTTGGGGATAATGTTGGAGCGATTGTTTACATGGG 1684  
Db 1621 ATATATGGCGGTATTTAGCACAGCCTTGGGGATAATGTTGGAGCGATTGTTTACATGGG 1680  
QY 1685 AACAAAGTGCCTTTGTCGAAAAATCTATACTAATGTGAAAAATTTGACTAGAGACCCAAAGAA 1744  
Db 1681 AACAAAGTGCCTTTGTCGAAAAATCTATACTAATGTGAAAAATTTGACTAGAGACCCAAAGAA 1740  
QY 1745 AACCTGSAACCTTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGACACAGCAAAA 1800  
Db 1741 AACCTGSAACCTTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGACACAGCAAAA 1796  
  
RESULT 4  
AAH46951  
ID AAH46951 standard; cDNA; 3076 BP.  
XX  
AC AAH46951;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Human secreted protein encoding cDNA (clone Id HBOEG11).  
DE  
DE Secreted protein; immunosuppressive; antiafthritic; antirheumatic;  
KW antiproliferative; cytosstatic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virucide; fungicide; human;  
KW opthalmological; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200155430-A1.  
XX  
PD 02-AUG-2001.  
XX  
PE 17-JAN-2001: 2001WO-US01431.  
XX  
PR 31-JAN-2000: 2000US-0179065.  
PR 04-FEB-2000: 2000US-0180628.  
PR 12-SEP-2000: 2000US-0231968.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Komatsoulis GA, Baker KP, Birse CE, Soppet DR, Olsen HS;  
PI Moore PA, Wei P, Ebner R, Duan DR, Shi Y, Choi GH, Fiscella M;  
PI Ni J, Ruben SW, Barash SC;  
XX  
DR WPI: 2001-476220/51.  
DR P-PSDB; AAB85541.  
XX  
PT 17 isolated nucleic acid molecules encoding human secreted proteins,  
PT used to preventing, treating or ameliorating a medical condition -  
XX

Claim 1; Page 426-427; 482pp; English.

The invention provides novel human secreted proteins and polynucleotides encoding them. The secreted proteins can be expressed by standard recombinant methodology. The secreted proteins and polynucleotides are used to prevent, treat or ameliorate a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. They can also be used in diagnosing a pathological condition. The antibodies to the proteins can also be used in alleviating symptoms associated with the disorders and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays (ELISA). Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. The present sequence represents a human secreted protein encoding cDNA.

Sequence 3076 BP; 940 A; 487 C; 617 G; 1028 T; 4 other;

Query Match 97.4%; Score 1753; DB 22; Length 3076;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 1762; Conservative 3; Mismatches 18; Indels 0; Gaps 0;

PS	18	CTGCTCGTCTGCTGCCCGGACCCGGGCGGAGCAGCAGACACACGCTATCAAGATAAA	77
XX			
CC	13	CTGCAGGTACCGGTCGCGGAATTCCTCGGCTCGACSCAGCGGCGGCGTATCAAGATAAA	72
CC			
CC	78	GAGGAAGTTGCTCTATGGATGAATCTGTTGGGCCCTACCAATAATCGTCAAGAAACATAT	137
CC			
CC	73	GAGGAAGTTGCTCTATGGATGAATCTGTTGGGCCCTACCAATAATCGTCAAGAAACATAT	132
CC			
CC	138	AGTACTTTTTCACCTCCATCTCTGTGGGTCAAAAAAGTATCATGCTATACCATGAA	197
CC			
CC	133	AGTACTTTTTCACCTCCATCTCTGTGGGTCAAAAAAGTATCATGCTATACCATGAA	192
CC			
CC	198	ACTCTGGGAGAACCTTCAAGGGGTTGAATGGAAATTTAGTGGCTGGATATTAATTT	257
CC			
CC	193	ACTCTGGGAGAACCTTCAAGGGGTTGAATGGAAATTTAGTGGCTGGATATTAATTT	252
CC			
CC	258	AAAGATGATGTGATGCCAGCCACTTACTCTGGAATTTAGATAAGAAAGAGAGAT	317
CC			
CC	253	AAAGATGATGTGATGCCAGCCACTTACTCTGGAATTTAGATAAGAAAGAGAGAT	312
CC			
CC	318	GCATTTGTATATGCCATAAAAAATCATTTACTGGTACCAGATGTACATAGATGATTTACCA	377
CC			
CC	313	GCATTTGTATATGCCATAAAAAATCATTTACTGGTACCAGATGTACATAGATGATTTACCA	372
CC			
CC	378	ATATGGGGTATTTGGTGGAGCTGATGAAATGGAGAGATTTACTATCTTTGGACCTAT	437
CC			
CC	373	ATATGGGGTATTTGGTGGAGCTGATGAAATGGAGAGATTTACTATCTTTGGACCTAT	432
CC			
CC	438	AAAAAATTTGAATAGTTTAAATGAAATCGAATTTGTTGATGTTAAATCTAAGTAGTAA	497
CC			
CC	433	AAAAAATTTGAATAGTTTAAATGAAATCGAATTTGTTGATGTTAAATCTAAGTAGTAA	492
CC			
CC	498	GGAAAGGTGAAACCTGGTTCCAAATCTAAAATCCAGATGTCATATTCAGTAAATGGAAA	557
CC			
CC	493	GGAAAGGTGAAACCTGGTTCCAAATCTAAAATCCAGATGTCATATTCAGTAAATGGAAA	552
CC			
CC	558	AGTCCAGATCTGAAATTTGAAGATCGATTTGACAAATATCTTGATCCGCTTTTTCAA	617
CC			
CC	553	AGTCCAGATCTGAAATTTGAAGATCGATTTGACAAATATCTTGATCCGCTTTTTCAA	612
CC			
CC	618	CATCGGATTCATTGGTTTTCATTTTCAACTCCTTCATCATGATGGTATCTCTTGGTGGC	677
CC			
CC	613	CATCGGATTCATTGGTTTTCATTTTCAACTCCTTCATCATGATGGTATCTCTTGGTGGC	672
CC			

QY	678	TTAGTTTCAATGATTTTAAATGAGAACATTAAGAAAAAGATTATGCTCGGTACAGTAAAGAG	737
DB			
DB	673	TTAGTTTCAATGATTTTAAATGAGAACATTAAGAAAAAGATTATGCTCGGTACAGTAAAGAG	732
QY	738	GAAGAAATGGATGATATGATGATAGAGACCTAGGAGATGAATATGGATGGAAACAGGTGCAT	797
DB			
DB	733	GAAGAAATGGATGATATGATGATAGAGACCTAGGAGATGAATATGGATGGAAACAGGTGCAT	792
QY	798	GGAGATGTTATTTAGACCATCAAGTCAACCCACCTGATATTTTCCCTCTGATTTGGTCTCGGA	857
DB			
DB	793	GGAGATGTTATTTAGACCATCAAGTCAACCCACCTGATATTTTCCCTCTGATTTGGTCTCGGA	852
QY	858	TGTCAGATATTTGCTGTCTCTCATCGTATATTTATTTGTCATATGATAGAAAGATTTATAT	917
DB			
DB	853	TGTCAGATATTTGCTGTCTCTCATCGTATATTTATTTGTCATATGATAGAAAGATTTATAT	912
QY	918	ACTGAGAGGGGATCAATGCTCAGTCAGTCAGGCATATTTGCTCTATGCTCTACGTCCTCAGTG	977
DB			
DB	913	ACTGAGAGGGGATCAATGCTCAGTCAGTCAGGCATATTTGCTCTATGCTCTACGTCCTCAGTG	972
QY	978	AATGGTTATTTGGAGGAAGTCTGTATGCTAGACAGGAGGAGAGATGGATATAAGACGAG	1037
DB			
DB	973	AATGGTTATTTGGAGGAAGTCTGTATGCTAGACAGGAGGAGAGATGGATATAAGACGAG	1032
QY	1038	ATGTTTATTTGGGGCATTCCTTATCCAGCTATGGTGTGGCAGCTCCCTTCTTCATCAAT	1097
DB			
DB	1033	ATGTTTATTTGGGGCATTCCTTATCCAGCTATGGTGTGGCAGCTCCCTTCTTCATCAAT	1092
QY	1098	TTTCATAGCCATTTATTTACCATGCTTCAAGAGCCATTCCTTTTGGAAACAATTTGGGCGGTT	1157
DB			
DB	1093	TTTCATAGCCATTTATTTACCATGCTTCAAGAGCCATTCCTTTTGGAAACAATTTGGGCGGTT	1152
QY	1158	TGTTGCACTGTTTTTTTCTGTTATTTCTCTCTAAATCTTTGGTACAATACTTTGGCCGA	1217
DB			
DB	1153	TGTTGCACTGTTTTTTTCTGTTATTTCTCTCTAAATCTTTGGTACAATACTTTGGCCGA	1212
QY	1218	AATCTGTCTAGGTCAAGCCCAACTTCTCTGCTGCTCAATGCTGTGCTCGTCTCTATACCG	1277
DB			
DB	1213	AATCTGTCTAGGTCAAGCCCAACTTCTCTGCTGCTCAATGCTGTGCTCGTCTCTATACCG	1272
QY	1278	GAGAAAAATTTGTTCAATGGAGCCCTGGGTTATTTGTTGCTGGTGGGAATTTTACCTTTT	1337
DB			
DB	1273	GAGAAAAATTTGTTCAATGGAGCCCTGGGTTATTTGTTGCTGGTGGGAATTTTACCTTTT	1332
QY	1338	GGTTCAATCTTTATTTGAAATGATTTTCACTCTTCACGCTCTTTCTGGGCATATAAGATCTAT	1397
DB			
DB	1333	GGTTCAATCTTTATTTGAAATGATTTTCACTCTTCACGCTCTTTCTGGGCATATAAGATCTAT	1392
QY	1398	TATGCTTATGGCTTTCATGATGCTGGTGGTGGTTATCCTGTGCTGCTGCTGCTGCTGCTG	1457
DB			
DB	1393	TATGCTTATGGCTTTCATGATGCTGGTGGTGGTTATCCTGTGCTGCTGCTGCTGCTGCTG	1452
QY	1458	ACTATTTGCTGCACATATTTTCTACTAAATGAGAAAGATTACCGGTGGCAATGGACAAGT	1517
DB			
DB	1453	ACTATTTGCTGCACATATTTTCTACTAAATGAGAAAGATTACCGGTGGCAATGGACAAGT	1512
QY	1518	TTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATGCTATTTCTTTTACTACTATTTT	1577
DB			
DB	1513	TTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATGCTATTTCTTTTACTACTATTTT	1572
QY	1578	TTTCAAAACAAAGATGATGGCTTATTTTCAACATCATTTTACTTTTGGATATATGGCGGTA	1637
DB			
DB	1573	TTTCAAAACAAAGATGATGGCTTATTTTCAACATCATTTTACTTTTGGATATATGGCGGTA	1632
QY	1638	TTTACACAGCCCTTGGGATAATGCTGAGGAGATTGGTTTACATGGGACACAGTGCCTTTT	1697
DB			
DB	1633	TTTACACAGCCCTTGGGATAATGCTGAGGAGATTGGTTTACATGGGACACAGTGCCTTTT	1692
QY	1698	GTCCGAAAAATCTACTACTAATGTTGAAAAATTGACTAGAGACCCCAAGAAACCTTGGAACTTT	1757
DB			
DB	1693	GTCCGAAAAATCTACTACTAATGTTGAAAAATTGACTAGAGACCCCAAGAAACCTTGGAACTTT	1752

Qy 1758 GGATCAATTTCTTTTCATAGGGTGAACCTGCACAGCAAAA 1800  
Db 1753 GGATCAATTTCTTTTCATAGGGTGAACCTGCACAGCAAAA 1795

RESULT 5  
AAF64216  
ID AAF64216 standard; cDNA; 2536 BP.

XX AAF64216;  
XX  
XX  
XX 06-APR-2001 (first entry)  
DE Human secreted protein gene 41 SEQ ID NO:51.

XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
XX antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
XX neotropic; neuroprotective; antibacterial; virucide; fungicide;  
XX ophthalmological; vulnerary; autoimmune disease; cardiovascular disorder;  
XX hyperproliferative disorders; cerebrovascular disorder; wound healing;  
XX nervous system disorder; ocular disorder; skin ageing; chemotaxis;  
XX food additive; ss.

XX Homo sapiens.  
OS  
XX WO200077026-A1.  
XX  
XX 21-DEC-2000.

XX 01-JUN-2000; 2000WO-US14973.  
XX  
XX 11-JUN-1999; 99US-0138630.  
XX (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.

XX Rosen CA, Ruben SM, Komatsoulis GA;  
PI WPI: 2001-071258/08.  
XX P-PSDB; AAB75546.  
XX

XX Nucleic acid molecules encoding human secreted proteins, used in  
XX preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
XX Parkinson's diseases and cancers -

XX Claim 1; Page 462-463; 542pp; English.

XX Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide  
XX sequences AAF64176 - AAF64224. The specification includes amino acid  
XX sequences AAB75555 - AAB75606 which represent fragments of the human  
XX secreted proteins, and protein sequences with which they share homology.  
XX The proteins and polynucleotides, their agonists and antagonists have  
XX activities dependent on the tissues and cells in which they are  
XX expressed, examples of these activities include, immunosuppressive;  
XX antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
XX vasotropic; cerebroprotective; neotropic; neuroprotective; antibacterial;  
XX virucide; fungicide; ophthalmological; and vulnerary. The proteins,  
XX polynucleotides, agonists and antagonists can be used to treat or detect  
XX or diagnose various diseases and disorders including, autoimmune  
XX diseases e.g. rheumatoid arthritis, hyperproliferative disorders  
XX e.g. neoplasms of the breast or liver, cardiovascular disorders  
XX e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
XX angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
XX infections caused by bacteria, viruses and fungi and ocular disorders  
XX e.g. corneal infection. The polypeptides can also be used to aid wound  
XX healing and epithelial cell proliferation, to prevent skin ageing due to  
XX sunburn, to maintain organs before transplantation, for supporting cell  
XX culture of primary tissues, to regenerate tissues and in chemotaxis. The  
XX polypeptides can also be used as a food additive or preservative to  
XX increase or decrease storage capabilities. Included in the invention are  
XX polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which  
XX are used in the isolation, identification and characterisation of the  
XX proteins of the invention.

Db	964	AGGAGGAGGAGATGGATATAAGCAGATGTTTATTGGGGCATTTCCCTATTACCCAGCATATGGT	1023
Qy	1073	GTGTGGCACTGCCCTTCTTCATCAATTTCAATAGCCATTTATTACCATGCTTCAAGAGCCAT	1132
Db	1024	GTGTGGCACTGCCCTTCTTCATCAATTTCAATAGCCATTTATTACCATGCTTCAAGAGCCAT	1093
Qy	1133	TCCTTTTGGAAACAATGGTGGCCGCTTTGTTGGCATCTGTTTTTTTTTTGTTATCTTCCCTCTAAA	1192
Db	1084	TCCTTTTGGAAACAATGGTGGCCGCTTTGTTGGCATCTGTTTTTTTTTTGTTATCTTCCCTCTAAA	1143
Qy	1193	TCCTTTTGGTACAAATACTTGGCCGGAATCTGTCAAGTCAAGCCAACTTTCCCTTGTGCTGT	1252
Db	1144	TCCTTTTGGTACAAATACTTGGCCGGAATCTGTCAAGTCAAGCCAACTTTCCCTTGTGCTGT	1203
Qy	1253	CAATGCTGTGCCCTGCTCCTATACCGGAGAAAAATGGTTTCATGGAGCCTGCGGTTATTGT	1312
Db	1204	CAATGCTGTGCCCTGCTCCTATACCGGAGAAAAATGGTTTCATGGAGCCTGCGGTTATTGT	1263
Qy	1313	TTGCGCTGGGTGGAAATTTTACCTTTTGGTTCAAATCTTTATTGAAATGTATTTTCATCTTCAC	1372
Db	1264	TTGCGCTGGGTGGAAATTTTACCTTTTGGTTCAAATCTTTATTGAAATGTATTTTCATCTTCAC	1323
Qy	1373	GTCTTCTGGGCATATAAGATCTATTATGTCTATGGCTTCATGATGCTGTGCTGGTTAT	1432
Db	1324	GTCTTCTGGGCATATAAGATCTATTATGTCTATGGCTTCATGATGCTGTGCTGGTTAT	1383
Qy	1433	CCCTGTGATTTGTGACTGCTCTGTGACTATTTGTGTGCACATATTTTCTACATAATGCAGA	1492
Db	1384	CCCTGTGATTTGTGACTGCTCTGTGACTATTTGTGTGCACATATTTTCTACATAATGCAGA	1443
Qy	1493	AGATTACCGGTGGCAATGGACAAAGTTTTCTCTCTGCTGCATCACTGCAATCTATGTTTA	1552
Db	1444	AGATTACCGGTGGCAATGGACAAAGTTTTCTCTCTGCTGCATCACTGCAATCTATGTTTA	1503
Qy	1553	CATGATATCCTTTTACTACTATATTTTTCAAAACAAGATGATGGCTTATTTCAAACATC	1612
Db	1504	CATGATATCCTTTTACTACTATATTTTTCAAAACAAGATGATGGCTTATTTCAAACATC	1563
Qy	1613	ATTTTACTTTTGGATATATGCGGTATTTAGCACAGCCTTGGGGATAAATGCTGGAGCGAT	1672
Db	1564	ATTTTACTTTTGGATATATGCGGTATTTAGCACAGCCTTGGGGATAAATGCTGGAGCGAT	1623
Qy	1673	TGGTTACATGGGAACAAGTGCCTTTTCTCCGAAAAATCTATACTAATGTGAAAAATTGACTA	1732
Db	1624	TGGTTACATGGGAACAAGTGCCTTTTCTCCGAAAAATCTATACTAATGTGAAAAATTGACTA	1683
Qy	1733	GAGACCCAGAAAAACCTGGAACCTTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGA	1792
Db	1684	GAGACCCAGAAAAACCTGGAACCTTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGA	1743
Qy	1793	CAGCAAAA 1800	
Db	1744	CAGCAAAA 1751	
RESULT 6			
AAH16543			
ID: AAH16543 standard; cDNA; 2459 BP.			
XX AAH16543;			
XX AC			
DT			
XX			
DE 26-JUN-2001 (first entry)			
Human cDNA sequence SEQ ID NO:15599.			
Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.			
XX Homo sapiens.			
OS			
PN EPI074617-A2.			
XX			
07-FEB-2001.			
XX			

PF	28-JUL-2000; 2000EP-0116126.	
XX		
PR	29-JUL-1999; 99JP-0248036.	
PR	27-AUG-1999; 99JP-0300253.	
PR	11-JAN-2000; 2000JP-0118776.	
PR	02-MAY-2000; 2000JP-0183767.	
PR	09-JUN-2000; 2000JP-0241899.	
XX		
PA	(HELI-) HELIX RES INST.	
XX		
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
XX		
DR	WPI; 2001-318749/34.	
XX		
PT	Primer sets for synthesizing polynucleotides, particularly the 5602	
PT	full-length cDNAs defined in the specification, and for the detection	
PT	and/or diagnosis of the abnormality of the proteins encoded by the	
PT	full-length cDNAs -	
XX		
PS	Claim 8; SEQ ID 15599; 2537pp + CD ROM; English.	
XX		
CC	The present invention describes primer sets for synthesizing 5602	
CC	full-length cDNAs defined in the specification. Where a primer set	
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary	
CC	to the complementary strand of a polynucleotide which comprises one of	
CC	the 5602 nucleotide sequences defined in the specification, where the	
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC	of an oligonucleotide comprising a sequence complementary to the	
CC	complementary strand of a polynucleotide which comprises a 5'-end	
CC	sequence and an oligonucleotide comprising a sequence complementary to a	
CC	polynucleotide which comprises a 3'-end sequence, where the	
CC	oligonucleotide comprises at least 15 nucleotides and the combination of	
CC	the 5'-end sequence/3'-end sequence is selected from those defined in	
CC	the specification. The primer sets can be used in antisense therapy and	
CC	in gene therapy. The primers are useful for synthesizing polynucleotides,	
CC	particularly full-length cDNAs. The primers are also useful for the	
CC	detection and/or diagnosis of the abnormality of the proteins encoded by	
CC	the full-length cDNAs. The primers allow obtaining of the full-length	
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and	
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB992446 to	
CC	AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632	
CC	represent oligonucleotides, all of which are used in the exemplification	
CC	of the present invention.	
XX		
SQ	Sequence 2459 BP; 728 A; 399 C; 490 G; 842 T; 0 other;	
	Query Match 94.9%; Score 1708.4; DB 22; Length 2459;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1709; Conservative 0; Mismatches 1; Indels 0; Gaps 0	
QY	91 TATGGATGAATACTGTGGGCCCTACCATAATCGTCAAGAACATATAAGTCTTTTCAC 150	
Db	1 TATGGATGAATACTGTGGGCCCTACCATAATCGTCAAGAACATATAAGTCTTTTCAC 60	
QY	151 TTCCATTCTGTGGGGTCAAAAAAGATCATCATCATCAACATCAACATCGGAGAAG 210	
Db	61 TTCCATTCTGTGGGGTCAAAAAAGATCATCATCATCAACATCGGAGAAG 120	
QY	211 CACTTCAAGGGGTGGAATTTAGTGGTCTGGATATTAATTTAAAGATGATGTGA 270	
Db	121 CACTTCAAGGGGTGGAATTTAGTGGTCTGGATATTAATTTAAAGATGATGTGA 180	
QY	271 TGCCAGCCACTTACTGTGAAATTTAGATATAAGAAAGAGAGATGCAATTTGTATATG 330	
Db	181 TGCCAGCCACTTACTGTGAAATTTAGATATAAGAAAGAGAGATGCAATTTGTATATG 240	
QY	331 CCATAAAAAATCATTTACTTGGTACCAGATGTACATAGATGATTTACCAGATGGGTATTG 390	
Db	241 CCATAAAAAATCATTTACTTGGTACCAGATGTACATAGATGATTTACCAGATGGGTATTG 300	
QY	391 TTGGTGGAGGCTGATGAAATGGAGAAGATTACTATCTTTGGACCTATAAAAACTTGAAA 450	

Db 301 TTGGTGGAGCTGTGAAATGGAGAGATTACTATCTTTTGGACCTATAAAAACTTGAAA 360  
QY 451 TAGGTTTTTAATGGAATCGAATCTTGATGTTAACTAAGTAGTGAAGGAAAGGTGAAC 510  
Db 361 TAGGTTTTTAATGGAATCGAATCTTGATGTTAACTAAGTAGTGAAGGAAAGGTGAAC 420  
QY 511 TGGTTCGAATACTAAATCCAGATGTCATATTCAGTAAATGGAAGGAAAGGTGCA 570  
Db 421 TGGTTCGAATACTAAATCCAGATGTCATATTCAGTAAATGGAAGGAAAGGTGCA 480  
QY 571 AATTGGAATCGATTGACAAATATCTTGATCGTCCCTTTTTCACATCGGATTCATT 630  
Db 481 AATTGGAATCGATTGACAAATATCTTGATCGTCCCTTTTTCACATCGGATTCATT 540  
QY 631 GGTGTTTCAATTTTCAACTCTTCATGATGTTGATCTTCTGGTGGGCTAGTTTCAATGA 690  
Db 541 GGTGTTTCAATTTTCAACTCTTCATGATGTTGATCTTCTGGTGGGCTAGTTTCAATGA 600  
QY 691 TTTTAATGAGAACAATTAGAAAGATATCTCGGTACAGTAAAGAGAAAGAAATGGATG 750  
Db 601 TTTTAATGAGAACAATTAGAAAGATATCTCGGTACAGTAAAGAGAAAGAAATGGATG 660  
QY 751 ATATGGATAGACCTAGGAGATGAATATGGATGGAACAGGTGCATGGAGATGATTTA 810  
Db 661 ATATGGATAGACCTAGGAGATGAATATGGATGGAACAGGTGCATGGAGATGATTTA 720  
QY 811 GACCATCAAGTCACCCACTGATATTTTCCCTCTGATTTGTTCTGGATGTCAGATATTG 870  
Db 721 GACCATCAAGTCACCCACTGATATTTTCCCTCTGATTTGTTCTGGATGTCAGATATTG 780  
QY 871 CTGTGCTCTCATCGTTTATTATTGTTGCAATGATAGAGATTTATATACCTAGAGGGGAT 930  
Db 781 CTGTGCTCTCATCGTTTATTATTGTTGCAATGATAGAGATTTATATACCTAGAGGGGAT 840  
QY 931 CAATGCTCAGTCAGCCATATTTGCTCTGATGTTGCTCAGTCCAGTCAATGGTTATTG 990  
Db 841 CAATGCTCAGTCAGCCATATTTGCTCTGATGTTGCTCAGTCCAGTCAATGGTTATTG 900  
QY 991 GAGGAAGTCTGTATGCTAGCAAGGAGAGAGATGATAAAGCAGATGTTTATTTGGG 1050  
Db 901 GAGGAAGTCTGTATGCTAGCAAGGAGAGAGATGATAAAGCAGATGTTTATTTGGG 960  
QY 1051 CATTCCTTATCCAGCTATGGTGTGGGACACTGCTCTTCATCAATTTTCATAGCCATTT 1110  
Db 961 CATTCCTTATCCAGCTATGGTGTGGGACACTGCTCTTCATCAATTTTCATAGCCATTT 1020  
QY 1111 ATTACCATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCGTTTGTGCATCTGTT 1170  
Db 1021 ATTACCATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCGTTTGTGCATCTGTT 1080  
QY 1171 TTTTGTGTTATTTCCCTCTAAATCTTCTGTTACAATACTTGGCGGAAATCTGTCAGGTC 1230  
Db 1081 TTTTGTGTTATTTCCCTCTAAATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1140  
QY 1231 AGCCCAACTTTCCCTGTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290  
Db 1141 AGCCCAACTTTCCCTGTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1291 TCATGGAGCCTCGGTTATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1350  
Db 1201 TCATGGAGCCTCGGTTATTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260  
QY 1351 TTGAATGTTATTTCACTTCCAGTCTTCTTGGGATATAAGATCTATTATGCTGATGGCT 1410  
Db 1261 TTGAATGTTATTTCACTTCCAGTCTTCTTGGGATATAAGATCTATTATGCTGATGGCT 1320  
QY 1411 TCATGATGCTGGTGGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1470  
Db 1321 TCATGATGCTGGTGGTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380  
QY 1471 CATATTTTCTACTAAATGCAAGATATACCGTGGCAATGGCAAGTTTCTCTGCTGCTG 1530  
Db 1381 CATATTTTCTACTAAATGCAAGATATACCGTGGCAATGGCAAGTTTCTCTGCTGCTG 1440

QY 1531 CATCAACTGCAATCTATGTTTACATGATTCCTTTTACTACTATTTTTCAAAAACAAGA 1590  
Db 1441 CATCAACTGCAATCTATGTTTACATGATTCCTTTTACTACTATTTTTCAAAAACAAGA 1500  
QY 1591 TGTATGCTTATTTTCAAAACATCATTTTACATTTTGGATATATGGCGGTATTTAGCACAGCCT 1650  
Db 1501 TGTATGCTTATTTTCAAAACATCATTTTACATTTTGGATATATGGCGGTATTTAGCACAGCCT 1560  
QY 1651 TGGGATAAATGCTGGAGCGATTGGTTACATGGGAACAAGTGCCTTTGTCGAAAAATCT 1710  
Db 1561 TGGGATAAATGCTGGAGCGATTGGTTACATGGGAACAAGTGCCTTTGTCGAAAAATCT 1620  
QY 1711 ATACTAATGTGAAAATTTGACTAGAGACCCCAAGAAAACTTGAACCTTTGGATCAATTTCTT 1770  
Db 1621 ATACTAATGTGAAAATTTGACTAGAGACCCCAAGAAAACTTGAACCTTTGGATCAATTTCTT 1680  
QY 1771 TTTTCATAGGGGTGGAACCTTGCACAGCAAAA 1800  
Db 1681 TTTTCATAGGGGTGGAACCTTGCACAGCAAAA 1710  
  
RESULT 7  
AAK94275  
ID AAK94275 standard; cDNA; 2864 BP.  
XX AAK94275;  
XX AC  
XX 06-NOV-2001 (first entry)  
DT  
DE Human full-length cDNA, SEQ ID NO: 2908.  
DE Human: full length cDNA; cDNA synthesis; oligo-capping; ss.  
KW Homo sapiens.  
XX OS  
XX EP1130094-A2.  
PN  
XX 05-SEP-2001.  
PD  
XX 07-JUL-2000; 2000EP-0114089.  
PF  
XX 08-JUL-1999; 99JP-0194486.  
PR 11-JAN-2000; 2000JP-0118774.  
PR 02-MAY-2000; 2000JP-0183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX  
DR WPI: 2001-524255/58.  
DR P-PSDB; AAM93355.  
XX  
XX 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation.  
PT  
XX  
XX Claim 8; SEQ ID NO 2908; 1380pp + sequence listing; English.  
XX  
CC The invention relates to primers for synthesising full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5' and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesising the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a full length  
CC human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX  
SQ Sequence 2864 BP; 834 A; 463 C; 569 G; 998 T; 0 other;



Query Match		74.6%	Score 1343.4;	DB 22;	Length 2864;		
Best Local Similarity		99.9%	Pred. No. 0;				
Matches 1344;		Conservative	0;	Mismatches	1;	Indels	0;
						Gaps	0;
QY	456	TTTTAAGGAATCGAATTTGGTGTGTTAACTCTAAGTGAAGAAAGGTGAAACTGGTT	515				
DB	1	TTTAATGGAATCGAATTTGGTGTGTTAACTCTAAGTGAAGAAAGGTGAAACTGGTT	60				
QY	516	CCAAATACTAAATCCAGATGTCATATTCAGTAAATGGAAGATCAGATGTGAATTT	575				
DB	61	CCAAATACTAAATCCAGATGTCATATTCAGTAAATGGAAGATCAGATGTGAATTT	120				
QY	576	GAAATCGGATTTGACAAATATCTTGATCCGTCCTTTTCAACATCGGATTCATTCGGTT	635				
DB	121	GAAATCGGATTTGACAAATATCTTGATCCGTCCTTTTCAACATCGGATTCATTCGGTT	180				
QY	636	TCAATTTTCAACTCCCTCATGATGGTGATCTTCCTTGGTGGGCTTAGTTCAATGATTTTA	695				
DB	181	TCAATTTTCAACTCCCTCATGATGGTGATCTTCCTTGGTGGGCTTAGTTCAATGATTTTA	240				
QY	696	ATCAGAACATTAAAGAAAGATTATGCTCGGTACAGTAAAGAGAAATGGATGATATG	755				
DB	241	ATCAGAACATTAAAGAAAGATTATGCTCGGTACAGTAAAGAGAAATGGATGATATG	300				
QY	756	GATAGAGACCTAGGAGATGAATATGGATGGAACAGGTGCATGGAGATGTATTTAGACCA	815				
DB	301	GATAGAGACCTAGGAGATGAATATGGATGGAACAGGTGCATGGAGATGTATTTAGACCA	360				
QY	816	TCAAGTACCCACTGATATTTTCTCTCTGATGGTCTGATGTCAGATATTTTGTGTG	875				
DB	361	TCAAGTACCCACTGATATTTTCTCTCTGATGGTCTGATGTCAGATATTTTGTGTG	420				
QY	876	TCCTCATCGTTATTTGTTGCAATGATAGAAATTTATATCTAGAGGGGATCAATG	935				
DB	421	TCCTCATCGTTATTTGTTGCAATGATAGAAATTTATATCTAGAGGGGATCAATG	480				
QY	936	CTCAGTACAGCCATTTTGTCTATGCTGCTACGTCCTCCAGTGAATGTTATTTGGAGGA	995				
DB	481	CTCAGTACAGCCATTTTGTCTATGCTGCTACGTCCTCCAGTGAATGTTATTTGGAGGA	540				
QY	996	AGTCTGTATGCTAGACAAGGAGGAGATGGATGAAGACAGATGTTATTTGGGGCATTC	1055				
DB	541	AGTCTGTATGCTAGACAAGGAGGAGATGGATGAAGACAGATGTTATTTGGGGCATTC	600				
QY	1056	CTTATCCACGCTATGTTGTGTCGACCTGCTCTTCTCATCAATTCATAGCCATTTATTAC	1115				
DB	601	CTTATCCACGCTATGTTGTGTCGACCTGCTCTTCTCATCAATTCATAGCCATTTATTAC	660				
QY	1116	CATGCTTCAAGAGCCATTTCTTTTGGAAACAATGGTGGCGTTTGTTCGATCTGTTTTTTT	1175				
DB	661	CATGCTTCAAGAGCCATTTCTTTTGGAAACAATGGTGGCGTTTGTTCGATCTGTTTTTTT	720				
QY	1176	GTTATTTCTCCTCTAAATCTGTTGGTGTACAACTCTTGGCGGAAATCTGTAGTCAGCC	1235				
DB	721	GTTATTTCTCCTCTAAATCTGTTGGTGTACAACTCTTGGCGGAAATCTGTAGTCAGCC	780				
QY	1236	ACTTTTCTGTCGTCATGCTGTGCTGCTGCTATACCGGAGAAAAATGGTTTCATG	1295				
DB	781	ACTTTTCTGTCGTCATGCTGTGCTGCTGCTATACCGGAGAAAAATGGTTTCATG	840				
QY	1296	GAGCCCTGCGGTTATTTCTTGGCTGGGTGGAATTTTACCTTTTGGTTCAATCTTTATTGAA	1355				
DB	841	GAGCCCTGCGGTTATTTCTTGGCTGGGTGGAATTTTACCTTTTGGTTCAATCTTTATTGAA	900				
QY	1356	ATGATTTTCAATCTACGCTCTTCTGGGCATATAAGATCTATATGTCATGGCTTCATG	1415				
DB	901	ATGATTTTCAATCTACGCTCTTCTGGGCATATAAGATCTATATGTCATGGCTTCATG	960				
QY	1416	ATGCTGGTGGTTATCTCTGTCATTTGACTGTCTGTGCTACTATTGTGTGCACATAT	1475				
DB	961	ATGCTGGTGGTTATCTCTGTCATTTGACTGTCTGTGCTACTATTGTGTGCACATAT	1020				

QY	1476	TTTCTACTAAATGACAGAGATTACCGTGGCAATGACAAAGTTTCTCTCTGCTGCATCA	1535
DB	1021	TTTCTACTAAATGACAGAGATTACCGTGGCAATGACAAAGTTTCTCTCTGCTGCATCA	1080
QY	1536	ACTGCAATCTATGTTTACATGATTCCTTTTACTACTATTTTTCACAAACAAGATGTAT	1595
DB	1081	ACTGCAATCTATGTTTACATGATTCCTTTTACTACTATTTTTCACAAACAAGATGTAT	1140
QY	1596	GGCTTATTTCAACATCATTTTACTTTGGATATATGCGGTATTTAGCACAGCCTTGGGG	1655
DB	1141	GGCTTATTTCAACATCATTTTACTTTGGATATATGCGGTATTTAGCACAGCCTTGGGG	1200
QY	1656	ATAATGTGTGGAGCGATTGTTACATGGGAACAAGTGCCTTTGTCGAAAAATCTATACT	1715
DB	1201	ATAATGTGTGGAGCGATTGTTACATGGGAACAAGTGCCTTTGTCGAAAAATCTATACT	1260
QY	1716	AATGTGAAATTTGACTAGAGACCCCAAGAAACCTTGAACCTTTGGATCAATTTCTTTTCA	1775
DB	1261	AATGTGAAATTTGACTAGAGACCCCAAGAAACCTTGAACCTTTGGATCAATTTCTTTTCA	1320
QY	1776	TAGGGGTGGAACCTTGCACAGCAAAA	1800
DB	1321	TAGGGGTGGAACCTTGCACAGCAAAA	1345
RESULT 8			
AAV28120			
ID	AAV28120 standard: cDNA; 1317 BP.		
XX	AC AAV28120;		
XX	DT 25-SEP-1998 (first entry)		
DE	Nucleotide sequence encoding non-adrenergic SM-binding protein.		
KW	Non-adrenergic SM binding protein; human; muscle; receptor; ss;		
XX	Iodocyanopindolol.		
OS	Homo sapiens.		
FH	Key Location/Qualifiers		
FT	CDS 1..1317		
FT	/*tag= a		
FT	/product= "non-adrenergic SM-binding protein"		
XX	EP848059-A1.		
XX	17-JUN-1998.		
XX	12-DEC-1996; 96EP-0402719.		
XX	12-DEC-1996; 96EP-0402719.		
XX	(VETI-) VETIGEN.		
XX	Lenzen G, Morooka S, Strosberg AD, Sugasawa T;		
XX	WPI: 1998-314474/28.		
DR	P-PSDB; AAW61371.		
XX	New iodocyanopindolol receptor polypeptide - and corresponding DNA,		
PT	antibodies, etc.		
XX	Claim 4; Page 22-23; 56pp; English.		
CC	The non-adrenergic SM binding protein is found in human muscles. It is		
CC	a member of a novel receptor class which have a receptor activity other		
CC	than that of beta3-adrenergic receptors. The protein comprises sites		
CC	which, when exposed on the cell surface, are capable of binding		
CC	iodocyanopindolol (ICYP) under blockade of alpha , beta 1, beta 2 and		
CC	beta 3 adrenergic receptors and serotonin 5-HT1A and 5-HT1B receptors.		
CC	The binding is saturable and reversible by displacement with the		
CC	stereoselective beta-adrenergic receptor agonist SM-11044 but not		



CC	isoproterenol, epinephrine, norepinephrine, serotonin, dopamine or BRL-37344.	
XX		
SQ	Sequence 1317 BP; 355 A; 211 C; 288 G; 463 T; 0 other;	
	Query Match 73.2%; Score 1317; DB 19; Length 1317;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 1317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	357 ATGTACATAGATATTTACCAATATGGGTATTTGGTGGAGCTGATGAAATGGAGAA	416
Db		
QY	1 ATGTACATAGATATTTACCAATATGGGTATTTGGTGGAGCTGATGAAATGGAGAA	60
Db		
QY	417 GATTACTATCTTTGGACCTATAAAAACTTGAATAGTGTTTAATGAAATCGAAATGTT	476
Db		
QY	61 GATTACTATCTTTGGACCTATAAAAACTTGAATAGTGTTTAATGAAATCGAAATGTT	120
QY	477 GATGTTAATCTAAGTGAAGGAAAGTGAACCTGGTTCCTCAATACTAAAATCCAGATG	536
Db		
QY	121 GATGTTAATCTAAGTGAAGGAAAGTGAACCTGGTTCCTCAATACTAAAATCCAGATG	180
QY	537 TCATATTTCAGTAAATGGAAGAGTCAAGATCTGAAATTTTCAAGATCGATTGACAAATAT	596
Db		
QY	181 TCATATTTCAGTAAATGGAAGAGTCAAGATCTGAAATTTTCAAGATCGATTGACAAATAT	240
QY	597 CTTGATCCGCTCTTTTCAACATCGGATTCATTTGGTGTTCATTTTCAACTCCTTCATG	656
Db		
QY	241 CTTGATCCGCTCTTTTCAACATCGGATTCATTTGGTGTTCATTTTCAACTCCTTCATG	300
QY	657 ATGGTATCTCTTGGTGGCTTAGTTTCAATGATTTTAATGAGAACATTAAGAAAGAT	716
Db		
QY	301 ATGGTATCTCTTGGTGGCTTAGTTTCAATGATTTTAATGAGAACATTAAGAAAGAT	360
QY	717 TATGCTCGGTACAGTAAAGAGGAGAAATGGATGATATGGATAGAGACCTAGAGATGAA	776
Db		
QY	361 TATGCTCGGTACAGTAAAGAGGAGAAATGGATGATATGGATAGAGACCTAGAGATGAA	420
QY	777 TATGGATGGAACAGGTGTCATGAGATGATTTAGACCATCAAGTACCACCTGATATTT	836
Db		
QY	421 TATGGATGGAACAGGTGTCATGAGATGATTTAGACCATCAAGTACCACCTGATATTT	480
QY	837 TCCTCTCTGATGGTCTGAGATGTCAGATATTTGCTGTCTCATCGTATATTATGTT	896
Db		
QY	481 TCCTCTCTGATGGTCTGAGATGTCAGATATTTGCTGTCTCATCGTATATTATGTT	540
QY	897 GCAATGATAGAGATTTATATCTAGAGGGGATCAATGCTCAGTACAGCCATATTGTC	956
Db		
QY	541 GCAATGATAGAGATTTATATCTAGAGGGGATCAATGCTCAGTACAGCCATATTGTC	600
QY	957 TATGCTGCTACGCTCCAGTGAATGGTATTTTGGAGGAAGTCTGTATGCTAGACAAGGA	1016
Db		
QY	601 TATGCTGCTACGCTCCAGTGAATGGTATTTTGGAGGAAGTCTGTATGCTAGACAAGGA	660
QY	1017 GGAAGGAGATGATATAAGCAGATGTTTATTTGGGGCATTCCTTATCCAGCATATGGTGTG	1076
Db		
QY	661 GGAAGGAGATGATATAAGCAGATGTTTATTTGGGGCATTCCTTATCCAGCATATGGTGTG	720
QY	1077 GGCACCTGCTCTTCATCAATTTTCATAGCCATTTATACCATGCTTCAAGAGCCATTCCT	1136
Db		
QY	721 GGCACCTGCTCTTCATCAATTTTCATAGCCATTTATACCATGCTTCAAGAGCCATTCCT	780
QY	1137 TTTTGGAAACATGTTGGCCGTTTCTGTCATCTGTTTTTTTGTATTTCTCTCAAAATCTTT	1196
Db		
QY	781 TTTTGGAAACATGTTGGCCGTTTCTGTCATCTGTTTTTTTGTATTTCTCTCAAAATCTTT	840
QY	1197 GTTGGTACAATACTTGGCCGAAATCTGTACAGTCAAGCCCACTTCTTGTCTGTGCAAT	1256
Db		
QY	841 GTTGGTACAATACTTGGCCGAAATCTGTACAGTCAAGCCCACTTCTTGTCTGTGCAAT	900
QY	1257 GCTGTGCTGCTCTTATACCGGAGAAAAATGTTTCATGGAGCCTCGGGTTATGTTTGC	1316
Db		
QY	901 GCTGTGCTGCTCTTATACCGGAGAAAAATGTTTCATGGAGCCTCGGGTTATGTTTGC	960

QY	1317 CTGGTGGAAATTTACCTTTTGGTTCAAATCTTATTTGAAATGATTTTCATCTTCAGCTCT	1376
Db		
QY	961 CTGGTGGAAATTTACCTTTTGGTTCAAATCTTATTTGAAATGATTTTCATCTTCAGCTCT	1020
QY	1377 TTCTGGGCATATAAGATCTATTATGCTCTATGGCTTCATGATGCTGGTGGTTCCTG	1436
Db		
QY	1021 TTCTGGGCATATAAGATCTATTATGCTCTATGGCTTCATGATGCTGGTGGTTCCTG	1080
QY	1437 TGCATTGTGACTCTGCTGTGACTATTGTGTGCACATATTTTCTACTAAATGCAGAAGAT	1496
Db		
QY	1081 TGCATTGTGACTCTGCTGTGACTATTGTGTGCACATATTTTCTACTAAATGCAGAAGAT	1140
QY	1497 TACCGTGGCAATGGACAAGTTTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATG	1556
Db		
QY	1141 TACCGTGGCAATGGACAAGTTTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATG	1200
QY	1557 TATTCCTTTTACTACTATTTTTCAAAACAAAGATGATATGGCTTATTTCAAACATCATTT	1616
Db		
QY	1201 TATTCCTTTTACTACTATTTTTCAAAACAAAGATGATATGGCTTATTTCAAACATCATTT	1260
QY	1617 TACTTTGGATATATGCGGTATTTAGCACAGCCTTGGGGATAATGTTGCGAGCGATT	1673
Db		
QY	1261 TACTTTGGATATATGCGGTATTTAGCACAGCCTTGGGGATAATGTTGCGAGCGATT	1317
DE	Human cDNA sequence SEQ ID NO:11070.	
OS	Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.	
OS	Homo sapiens.	
PN	EP1074617-A2.	
PD	07-FEB-2001.	
PF	28-JUL-2000; 2000EP-0116126.	
PR	29-JUL-1999; 99JP-0248036.	
PR	27-AUG-1999; 99JP-0300253.	
PR	11-JAN-2000; 2000JP-0118776.	
PR	02-MAY-2000; 2000JP-0183767.	
PR	09-JUN-2000; 2000JP-0241899.	
PA	(HELI-) HELIX RES INST.	
PI	Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;	
PI	Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;	
WPI	2001-318749/34.	
PT	Primer sets for synthesizing polynucleotides, particularly the 5602	
PT	full-length cDNAs defined in the specification, and for the detection	
PT	and/or diagnosis of the abnormality of the proteins encoded by the	
PT	full-length cDNAs.	
PS	Claim 8; SEQ ID 11070; 2537pp + CD ROM; English.	
CC	The present invention describes primer sets for synthesizing 5602	
CC	full-length cDNAs defined in the specification. Where a primer set	
CC	comprises: (a) an oligo-dT primer and an oligonucleotide complementary	
CC	to the complementary strand of a polynucleotide which comprises one of	
CC	the 5602 nucleotide sequences defined in the specification, where the	
CC	oligonucleotide comprises at least 15 nucleotides; or (b) a combination	
CC	of an oligonucleotide comprising a sequence complementary to the	
CC	complementary strand of a polynucleotide which comprises a 5'-end	

CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX  
SQ Sequence 2311 BP; 655 A; 382 C; 458 G; 816 T; 0 other;

Query Match 61.3%; Score 1104; DB 22; Length 2311;  
Best Local Similarity 100.0%; Pred. No. 3.3e-276;  
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 697 TGAGAACATTAGAAAGATTATGCTCGCTACAGTAAGACGAAGAAATCGATATGG 756  
DB 1 TGAGAACATTAGAAAGATTATGCTCGCTACAGTAAGACGAAGAAATCGATATGG 60  
QY 757 ATAGAGACCTAGGAGATGAATATGGATGAAACAGCTGCATGGAGATGATTTAGACCAT 816  
DB 61 ATAGAGACCTAGGAGATGAATATGGATGAAACAGCTGCATGGAGATGATTTAGACCAT 120  
QY 817 CAAGTCACCCACTGATATTTTCCCTCTCTGATTGGTTCAGATGTCAGATATTTGCTGTG 876  
DB 121 CAAGTCACCCACTGATATTTTCCCTCTCTGATTGGTTCAGATGTCAGATATTTGCTGTG 180  
QY 877 CTCTCATCTGTTATTTGTTGCAATCATAGAGATTTATATCTAGAGGGGATCAATGC 936  
DB 181 CTCTCATCTGTTATTTGTTGCAATCATAGAGATTTATATCTAGAGGGGATCAATGC 240  
QY 937 TCAGTACAGCCATATTTGTTCTATGCTGTACCTGCCAGTGAATGGTATTTTGGAGGAA 996  
DB 241 TCAGTACAGCCATATTTGTTCTATGCTGTACCTGCCAGTGAATGGTATTTTGGAGGAA 300  
QY 997 GTCTGTATGCTAGACAGGAGGAGATGGATGAAGCAGATGTTTATTTGGGGCATTCC 1056  
DB 301 GTCTGTATGCTAGACAGGAGGAGATGGATGAAGCAGATGTTTATTTGGGGCATTCC 360  
QY 1057 TTATCCACGCTATGTTGCTGTCACCTGCCCTCTCATCAATTTTATAGCCATTTATACC 1116  
DB 361 TTATCCACGCTATGTTGCTGTCACCTGCCCTCTCATCAATTTTATAGCCATTTATACC 420  
QY 1117 ATGCTTCAAGACCCATTCCTTTTGGAAACAATGGTGGCCGCTTTGTTGCATCTCTTTTGTG 1176  
DB 421 ATGCTTCAAGACCCATTCCTTTTGGAAACAATGGTGGCCGCTTTGTTGCATCTCTTTTGTG 480  
QY 1177 TTATTTCTCTCTAAATCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1236  
DB 481 TTATTTCTCTCTAAATCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540  
QY 1237 ACTTTTCCCTGCTGTCATGCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1296  
DB 541 ACTTTTCCCTGCTGTCATGCTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
QY 1297 AGCCTCGCGTTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356  
DB 601 AGCCTCGCGTTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 1357 TGTATTTTCATCTTACGCTCTTTTCTGGGATATAAGATCTATTTATGCTATGCTGCTGCTGCTGCTG 1416  
DB 661 TGTATTTTCATCTTACGCTCTTTTCTGGGATATAAGATCTATTTATGCTATGCTGCTGCTGCTGCTG 720  
QY 1417 TGCTG 1476  
DB 721 TGCTG

QY 1477 TTCTACTAAATGCAGAAGATTACCGTGGCAATGGCAAGTTTCTCTCTGCTGCTGATCAA 1536  
DB 781 TTCTACTAAATGCAGAAGATTACCGTGGCAATGGCAAGTTTCTCTCTGCTGCTGATCAA 840  
QY 1537 CTGCAATCTATGTTTACATGATGTTCTTTTACTACTATTTTTCAAAACAAAGATGATG 1596  
DB 841 CTGCAATCTATGTTTACATGATGTTCTTTTACTACTATTTTTCAAAACAAAGATGATG 900  
QY 1597 GCTTATTTCAACATCATTTTACTTTGGATATATGGCGTATTTTAGCACAGCTTTGGGGA 1656  
DB 901 GCTTATTTCAACATCATTTTACTTTGGATATATGGCGTATTTTAGCACAGCTTTGGGGA 960  
QY 1657 TAATGTGTGGAGCGATTTGGTTACATGGGAACAAGTGCCTTTGTCGGAATAATCTATCTA 1716  
DB 961 TAATGTGTGGAGCGATTTGGTTACATGGGAACAAGTGCCTTTGTCGGAATAATCTATCTA 1020  
QY 1717 ATGCTAAAATTTGACTAGAGACCCCAAGAAACCTGGAACCTTTGGATCAATTTCTTTTCAT 1776  
DB 1021 ATGCTAAAATTTGACTAGAGACCCCAAGAAACCTGGAACCTTTGGATCAATTTCTTTTCAT 1080  
QY 1777 AGGGGTGGAACCTTGACACAGCAAAA 1800  
DB 1081 AGGGGTGGAACCTTGACACAGCAAAA 1104  
RESULT 10  
AAK94454  
ID AAK94454 standard; cDNA; 2299 BP.  
XX AC AAK94454;  
XX XX  
XX DT 06-NOV-2001 (first entry)  
XX XX  
XX DE Human full-length cDNA, SEQ ID NO: 3256.  
XX XX  
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping; ss.  
XX OS Homo sapiens.  
XX OS  
XX PN EP1130094-A2.  
XX XX  
XX PD 05-SEP-2001.  
XX PF 07-JUL-2000; 2000EP-0114089.  
XX XX  
XX PR 08-JUL-1999; 99JP-0194486.  
XX PR 11-JAN-2000; 2000JP-0118774.  
XX PR 02-MAY-2000; 2000JP-0183765.  
XX XX  
XX PA (HELI-) HELIX RES INST.  
XX XX  
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX XX  
XX DR WPI: 2001-524255/58.  
XX DR P-PSDB: AAM93524.  
XX XX  
XX PT 830 Primers useful for synthesizing full length cDNA clones and their  
XX use in genetic manipulation -  
XX PS Claim 8; SEQ ID NO 3256; 1380pp + sequence listing; English.  
XX XX  
XX CC The invention relates to primers for synthesising full length cDNA  
XX CC clones. 830 cDNA molecules encoding a human protein have been  
XX CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
XX CC molecules have been determined. Primers for synthesising the full length  
XX CC cDNA are useful for clarifying the function of the protein encoded by  
XX CC the cDNA. The full length clones were obtained by construction of full  
XX CC length enriched cDNA libraries that were synthesised by the oligo-capping  
XX CC method. The primers enable the production of the full length cDNA easily  
XX CC without any special methods. The present sequence is a full length  
XX CC human cDNA of the invention.  
XX Note: The sequence data for this patent did not form part of the printed

```
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 2299 BP; 649 A; 380 C; 457 G; 813 T; 0 other;

Query Match          60.6%; Score 1090.4; DB 22; Length 2299;
Best Local Similarity 99.9%; Pred. No. 1.1e-272;
Matches 1091; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 709 GAAAGATTATGCTCGGTACAGTAAAGAGGAAGAAATGGATGATATGATAGAGACCTAG 768
DB 1 GAAAGATTATGCTCGGTACAGTAAAGAGGAAGAAATGGATGATATGATAGAGACCTAG 60

QY 769 GAGATGAATATGGATGGAACAGGTGCATGGAGATGATTTAGACCATCAAGTCACCCAC 828
DB 61 GAGATGAATATGGATGGAACAGGTGCATGGAGATGATTTAGACCATCAAGTCACCCAC 120

QY 829 TGATATTTTCTCTCTGATTTGGTTCGGATGTCAGATATTTGCTGTCTCTCATCGTTA 888
DB 121 TGATATTTTCTCTCTGATTTGGTTCGGATGTCAGATATTTGCTGTCTCTCATCGTTA 180

QY 889 TTATTTGTCGAATCATAGAGATTATATACTGAGAGGGATCAATGCTCAGTACAGCCA 948
DB 181 TTATTTGTCGAATCATAGAGATTATATACTGAGAGGGATCAATGCTCAGTACAGCCA 240

QY 949 TATTTGCTATGCTGTACGTCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTA 1008
DB 241 TATTTGCTATGCTGTACGTCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTA 300

QY 1009 GACAAGGAGGAAGAGATGATGAAGCAGATGTTATTTGGGCATTCCTTATCCCAGCTA 1068
DB 301 GACAAGGAGGAAGAGATGATGAAGCAGATGTTATTTGGGCATTCCTTATCCCAGCTA 360

QY 1069 TGTGTGTGGCACTGCCCTCTTCATCAATTTTCATAGCCATTTATTACCATGCTTCAAGAG 1128
DB 361 TGTGTGTGGCACTGCCCTCTTCATCAATTTTCATAGCCATTTATTACCATGCTTCAAGAG 420

QY 1129 CCATTCCTTTTGGAAACAATGGTGGCCGTTTGGTGCATCTGTTTTTGGTATTCTTCCTC 1188
DB 421 CCATTCCTTTTGGAAACAATGGTGGCCGTTTGGTGCATCTGTTTTTGGTATTCTTCCTC 480

QY 1189 TAAATCTTGTGTACAAATCTTGGCGGAATCTGTGTCAGTCAGCCCAACTTTCCTTGTG 1248
DB 481 TAAATCTTGTGTACAAATCTTGGCGGAATCTGTGTCAGTCAGCCCAACTTTCCTTGTG 540

QY 1249 GTGTCAATGCTGTGCCCTGCTTATACCGGAGAAAAATGGTTCATGAGCCCTGCGGTTA 1308
DB 541 GTGTCAATGCTGTGCCCTGCTTATACCGGAGAAAAATGGTTCATGAGCCCTGCGGTTA 600

QY 1309 TTGTTGCCCTGGGTGGAAATTTTACCTTTTGGTTCAATCTTTATTTGAATGATTTTCATCT 1368
DB 601 TTGTTGCCCTGGGTGGAAATTTTACCTTTTGGTTCAATCTTTATTTGAATGATTTTCATCT 660

QY 1369 TCAGGCTTTCTGGGCATATAAGATCTATTATGCTATGCTTCATGCTTCATGCTGGTCTG 1428
DB 661 TCAGGCTTTCTGGGCATATAAGATCTATTATGCTATGCTTCATGCTTCATGCTGGTCTG 720

QY 1429 TTATCTCTGTGCATTTGTGACTGTCTGTGACTATTGTGTGCACATATTTTCTACTAAATG 1488
DB 721 TTATCTCTGTGCATTTGTGACTGTCTGTGACTATTGTGTGCACATATTTTCTACTAAATG 780

QY 1489 CAGAAGATTACCGGTGGCAATGACAAAGTTTTCTCTGCTGTCATCAACTGCAATCTATG 1548
DB 781 CAGAAGATTACCGGTGGCAATGACAAAGTTTTCTCTGCTGTCATCAACTGCAATCTATG 840

QY 1549 TTTTACATGATTTCTTTTACTACTATTTTTCAAAACAAGATGATGCGTTTATTTCAA 1608
DB 841 TTTTACATGATTTCTTTTACTACTATTTTTCAAAACAAGATGATGCGTTTATTTCAA 900

QY 1609 CATCATTTTACTTTTGGATATATGCGGTATTTAGCACAGCCTTTGGGGAATATGTGGAG 1668
DB 901 CATCATTTTACTTTTGGATATATGCGGTATTTAGCACAGCCTTTGGGGAATATGTGGAG 960

QY 1669 CGATTGGTTACATGGGAACAAGTGCCTTTGTGCGGAAAAATCTATACTAATGTGAAAAATTG 1728
```

```
DB 961 CGATTGGTTACATGGGAACAAGTGCCTTTGTCCGAAAAATCTATACTAATGCTGAAAAATTG 1020
QY 1729 ACTAGACACCAAGAAAACCTTGGAACTTTGGATCAATTTCTTTTTCATAGGGTGGAACT 1788
DB 1021 ACTAGACACCAAGAAAACCTTGGAACTTTGGATCAATTTCTTTTTCATAGGGTGGAACT 1080
QY 1789 TGCACACGCAAAA 1800
DB 1081 TGCACACGCAAAA 1092

RESULT 11
AAV28121
ID AAV28121 standard; DNA; 965 BP.
XX
AC AAV28121:
XX
DT 25-SEP-1998 (first entry)
XX
DE Non-adrenergic SM-binding protein 900 bp probe.
XX
KW Non-adrenergic SM binding protein; human; muscle; receptor; ss; probe;
KW Iodocyanopindolol.
XX
OS Homo sapiens.
XX
PN EP848059-A1.
XX
PD 17-JUN-1998.
XX
PF 12-DEC-1996; 96BP-0402719.
XX
PR 12-DEC-1996; 96BP-0402719.
XX
PA (VETI-) VETIGEN.
XX
PI Lenzen G, Morooka S, Strosberg AB, Sugawara T;
XX WPI; 1998-314474/28.
XX
PT New iodocyanopindolol receptor polypeptide - and corresponding DNA,
PT antibodies, etc.
XX
PS Claim 5; Page 23-24; 56pp; English.
XX
CC The probes AAV28121-V28122 were used in the isolation of non-adrenergic
CC receptors from human skeletal muscle. This includes the non-adrenergic
CC SM binding protein which is a member of a novel receptor class which
CC have a receptor activity other than that of beta3-adrenergic receptors.
CC The protein comprises sites which, when exposed on the cell surface, are
CC capable of binding iodocyanopindolol (ICYP) under blockade of alpha ,
CC beta 1, beta 2 and beta 3 adrenergic receptors and serotonin 5-HT1A and
CC 5-HT1B receptors. The binding is saturable and reversible by
CC displacement with the stereoselective beta-adrenergic receptor agonist
CC SM-11044 but not isoproterenol, epinephrine, norepinephrine, serotonin,
CC dopamine or BRL-37344.
XX
SQ Sequence 965 BP; 246 A; 163 C; 214 G; 342 T; 0 other;

Query Match          53.5%; Score 963.4; DB 19; Length 965;
Best Local Similarity 99.9%; Pred. No. 7e-240;
Matches 964; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 531 CAGATGTCATATTTCAGTAAATGGAANAACTCAGATGTGAAATTTGAAGATCGATTGAC 590
DB 1 CAGATGTCATATTTCAGTAAATGGAANAACTCAGATGTGAAATTTGAAGATCGATTGAC 60
QY 591 AATATCTTCATCGCTCTTTTTCACATCGGATTCATTTGGTTTCAATTTTCAACTCC 650
DB 61 AATATCTTCATCGCTCTTTTTCACATCGGATTCATTTGGTTTCAATTTTCAACTCC 120
QY 651 TTCATGATGCTGATCTTCTTGGTGGGCTTAGTTTCAATGATTTTAAATGAGAACAATTAAGA 710
```



Db 241 |TCAGTACAGCCATATTGTTGCTATGCTGCTACGCTCCAGTGAATGTTATTTGGAGGAA 300  
Qy 997 |GTCGTATGCTAGACAAGGAGGAGAGATGGATAAAGCAGATGTTTATTGGGGCATTC 1056  
Db 301 |GTCGTATGCTAGACAAGGAGGAGAGATGGATAAAGCAGATGTTTATTGGGGCATTC 360  
Qy 1057 |TATCCAGCTATGTTGCTGTCGACCTGCTTCATCAATTTATAGCCATTATTACC 1116  
Db 361 |TATCCAGCTATGTTGCTGTCGACCTGCTTCATCAATTTATAGCCATTATTACC 420  
Qy 1117 |ATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCCGTTGTTGCATCTGTTTTTG 1176  
Db 421 |ATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCCGTTGTTGCATCTGTTTTTG 480  
Qy 1177 |TATTCCTTCTCTAAATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1236  
Db 481 |NTATCTTCTCTAAATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540  
Qy 1237 |ACTTTCCTTGTGCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1296  
Db 541 |ACTTTCCTTGTGCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
Qy 1297 |AGCTCGCGTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1356  
Db 601 |AGCTCGCGTATTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
Qy 1357 |TGATTTTCATCT 1414  
Db 661 |TGATTTTCATCT 720  
Qy 1415 |GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1474  
Db 721 |GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780  
Qy 1475 |TT 1476  
Db 781 |TT 782

RESULT 13  
ABL90852  
ID ABL90852 standard; cDNA; 1070 BP.  
XX  
AC ABL90852;  
XX  
DT 24-MAY-2002 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 1414.  
XX  
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;  
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antilulcer;  
KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;  
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;  
KW neurological disease; infection; human; secreted protein; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200190304-A2.  
XX  
PD 29-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-US16450.  
XX  
PR 19-MAY-2000; 2000US-205515P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Birse CE, Rosen CA;  
XX  
DR WPI; 2002-122018/16.  
DR P-PSDB; ABB90443.  
XX

PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and  
PT prevention of neural, immune system, muscular, reproductive,  
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative  
PT disorders -  
XX  
PS Claim 4; SEQ ID NO 1414; 2081pp + Sequence Listing; English.  
XX  
CC The invention relates to novel genes (ABL89449-ABL90853) and proteins  
CC (ABB9040-ABB90444) useful for preventing, treating or ameliorating  
CC medical conditions e.g. by protein or gene therapy. The genes are  
CC isolated from a range of human tissues disclosed in the specification.  
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful  
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
CC haemolytic anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;  
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
CC and parasitic infections.  
CC Note: the sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1070 BP; 262 A; 206 C; 227 G; 363 T; 12 other;  
  
Query Match 39.8%; Score 716.4; DB 24; Length 1070;  
Best Local Similarity 92.6%; Pred. No. 9.9e-176;  
Matches 803; Conservative 0; Mismatches 11; Indels 53; Gaps 3;  
  
Qy 987 |TTTGGAGGAGTCTGATGCTAGACAAGGAGGAGATGGATAAAGCAGATGTTTATT 1046  
Db 20 |TTTGGAGGCTCTCTNATGAGACAAGGAGGAGACATGGATAAAGCAGATGTTTATT 79  
Qy 1047 |GGGGCATTCCTTATCCAGCTATGGTGTGGCACTGCCTTCTTCATCAATTTATAGCC 1106  
Db 80 |GGGGCATTCCTTATCCAGCTATGGTGTGGCACTGCCTTCTTCATCAATTTATAGCC 139  
Qy 1107 |ATTATTACATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCGGTTGTTGCATC 1166  
Db 140 |ATTATTACATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCGGTTGTTGCATC 199  
Qy 1167 |TGTTTTTTTGTATTCTCTCTAAATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1226  
Db 200 |TGTTTTTTTGTATTCTCTCTAAATCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 259  
Qy 1227 |GTCAGCCCAACTTTCCTTGTGCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1286  
Db 260 |GTCAGCCCAACTTTCCTTGTGCTGCTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319  
Qy 1287 |TGGTTTCATGGAGCTCGGTTATTGTTGCTGGGTTGGAATTTTACCTTTTGGTTCAATC 1346  
Db 320 |TGGTTTCATGGAGCTCGGTTATTGTTGCTGGGTTGGAATTTTACCTTTTGGTTCAATC 379  
Qy 1347 |TTTATTGAAATGATTTCATCTTCAGTCTTCTCTGGGATATAAGATCTATTATGCTAT 1406  
Db 380 |TTTATTGAAATGATTTCATCTTCAGTCTTCTCTGGGATATAAGATCTATTATGCTAT 439  
Qy 1407 |GGCTTCATGATGCTGGTGTGTTATCTGTCGATTTGACGTGTGCTGTGATATTG 1466  
Db 440 |GGCTTCATGATGCTGGTGTGTTATCTGTCGATTTGACGTGTGCTGTGATATTG 499  
Qy 1467 |TGCACATATTTTCTACTAAATGCAGAA--GATTACCGGT----- 1503  
Db 500 |TGCACATATTTTCTACTAAATGCAGNAAGNATACCGGTGTGCCATTCATTCAAAAGNAG 559  
Qy 1504 |-----GGCAATGGACAAGTTTCT 1534  
Db 560 |ATTATCT 619  
Qy 1535 |AACTGCAATCTATGTTTACATGTTATTCCTTTTACTACTATTTTTCAAAACAAGATGTA 1594



02-MAY-2000: 2000JP-0183765.

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
WPI; 2001-524255/58.

830 Primers useful for synthesizing full length cDNA clones and their  
use in genetic manipulation -

Example 11; SEQ ID NO 1864; 1380pp + sequence listing; English.

The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence was used as the representative sequence from a human clone which was used in homology searches to identify the clone.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from BPO.

Sequence 756 BP; 183 A; 135 C; 182 G; 251 T; 5 other;

Query Match 38.7%; Score 696.6; DB 22; Length 756;  
Best Local Similarity 99.3%; Pred. No. 1.2e-170;  
Matches 710; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 709 GAAAGATTATGCTCCGTACAGTAAAGAGCAAGAAATGGATGATGGATAGAGACCTAG 768

Db 1 GAAAGATTATGCTCCGTACAGTAAAGAGCAAGAAATGGATGATGGATAGAGACCTAG 60

QY 769 GAGATGAATATGATGGAACAGGTGCATGGAGATGATTTAGACCATCAAGTCACCCAC 828

Db 61 GAGATCAATATGGATGGAACAGGTGCATGGAGATGATTTAGACCATCAAGTCACCCAC 120

QY 829 TGATATTTTCCCTCTCGATTGGTTCGGATGTCAGATATTTGCTGTCTCTCATCGTTA 888

Db 121 TGATATTTTCCCTCTCGATTGGTTCGGATGTCAGATATTTGCTGTCTCTCATCGTTA 180

QY 889 TTATTCTTCAATGATAGAGATTTATATACTAGAGGGGATCAATGCTCAGTACAGCCA 948

Db 181 TTATTCTTCAATGATAGAGATTTATATACTAGAGGGGATCAATGCTCAGTACAGCCA 240

QY 949 TATTTGCTATGCTGCTAGCTCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTA 1008

Db 241 TATTTGCTATGCTGCTAGCTCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTA 300

QY 1009 GACAGGAGGAGGAGATGGATAAAGCAGATGTTTATCGGGCATTCTTATCCCCAGCTA 1068

Db 301 GACAAGGAGGAGGAGATGGATAAAGCAGATGTTTATCGGGCATTCTTATCCCCAGCTA 360

QY 1069 TGGTGTGGCCACTGCCTTCTTCATCAATTTTCATGCCAATTTATTACCATGCTTCAAGAG 1128

Db 361 TGGTGTGGCCACTGCCTTCTTCATCAATTTTCATGCCAATTTATTACCATGCTTCAAGAG 420

QY 1129 CCATTCCTTTTGGAAACAATGGTGGCGTTTGTTCATCTGTTTGTGTTATTTCTTCCTC 1188

Db 421 CCATTCCTTTTGGAAACAATGGTGGCGTTTGTTCATCTGTTTGTGTTATTTCTTCCTC 480

QY 1189 TAAATCTTGTGGTACAATACTTGGCCGAAATCTGTAGGTCAGCCCAACTTTTCCTTTGTC 1248

Db 481 TAAATCTTGTGGTACAATACTTGGCCGAAATCTGTAGGTCAGCCCAACTTTTCCTTTGTC 540

QY 1249 GTGCTCAATGCTGTGCTCTGCTATACCGGAGAAAAATGGTTCAATGGAGCCTGGGGTTA 1308

Db 541 GTGCTCAATGCTGTGCTCTGCTATACCGGAGAAAAATGGTTCAATGGAGCCTGGGGTTA 600

QY 1309 TTGTTTGGCTGGGTGGGAATTTTACCTTTTGGTTCAAATCTTTATTGAATGTATTTTCATCT 1368  
|||||  
Db 601 TTGTTTGGCTGGGTGGGAATTTTACCTTTTGGGTCAATCTTTATTGAATGTATTTTCATCT 660  
|||||  
QY 1369 TCACGCTTTTCTGGGCATATAAGATCTATTTATGT-CTATGGCTTCATGATGCTGG 1422  
|||||  
Db 661 TCACGCTTTTCTGGGCATATAAGATCTATTTAAGTCCAAAGGCTTCATGATGCTGG 715  
|||||

Search completed: January 20, 2003, 16:37:59  
Job time : 404 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 20, 2003, 07:18:30 ; Search time 4565 Seconds  
(without alignments)  
11475.367 Million cell updates/sec

Title: US-09-319-724A-13  
Perfect score: 1800  
Sequence: 1 ccgcgcgcgtgtgctgctg.....gtggaactgtcacagcaaaa 1800

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_htg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_on.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_htg\_hum.\*  
31: em\_htg\_inv.\*  
32: em\_htg\_other.\*  
33: em\_htg\_mus.\*  
34: em\_htg\_pln.\*  
35: em\_htg\_rod.\*  
36: em\_htg\_mam.\*  
37: em\_htg\_vrt.\*  
38: em\_sy.\*  
39: em\_htgo\_hum.\*  
40: em\_htgo\_mus.\*  
41: em\_htgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1800	100.0	1800	6	A91337	A91337 Sequence 13
2	1800	100.0	1804	9	AF116347	AF116347 Homo sapi
3	1798.4	99.9	3592	9	AF269150	AF269150 Homo sapi
4	1709	94.9	3389	9	AF160213	AF160213 Homo sapi
5	1708.4	94.9	2459	9	AK027788	AK027788 Homo sapi
6	1632.4	90.7	2870	9	BC020959	BC020959 Homo sapi
7	1551	86.2	3143	10	AF269151	AF269151 Mus muscu
8	1450.6	80.6	2973	9	AK000756	AK000756 Homo sapi
9	1386.2	77.0	2848	10	BC004799	BC004799 Mus muscu
10	1317	73.2	1317	6	A78983	A78983 Sequence 2
11	1317	73.2	1317	6	A91326	A91326 Sequence 2
12	1104	61.3	2311	9	AK027438	AK027438 Homo sapi
13	963.4	53.5	965	6	A78984	A78984 Sequence 3
14	963.4	53.5	965	6	A91327	A91327 Sequence 3
15	771.4	42.9	2662	3	AY071401	AY071401 Drosophil
c	631.6	35.1	189746	2	AC114360	AC114360 Rattus no
16	577.8	32.1	1888	8	AY081536	AY081536 Arabidops
17	577.8	32.1	2062	8	AY054666	AY054666 Arabidops
18	577.8	32.1	2094	8	AF446365	AF446365 Arabidops
19	577.8	32.1	2094	8	AY063907	AY063907 Arabidops
20	577.8	32.1	2113	8	AY063907	AY063907 Arabidops
21	563.6	31.3	84961	2	AC014810	AC014810 Drosophil
c	563.6	31.3	163115	3	AC010661	AC010661 Drosophil
22	563.6	31.3	182223	3	AC023674	AC023674 Drosophil
c	563.6	31.3	286583	3	AE003565	AE003565 Drosophil
23	500.2	27.8	1650	8	AY056242	AY056242 Arabidops
24	455.4	25.3	783	6	AX136534	AX136534 Sequence
25	455.4	25.3	1976	6	AX136235	AX136235 Sequence
26	409.4	22.7	418	6	AX071516	AX071516 Sequence
27	389.8	21.7	96413	10	AL831779	AL831779 Mouse DNA
c	355	19.7	180752	2	AC114392	AC114392 Rattus no
28	319.4	17.7	58069	3	AC024775	AC024775 Caenorhab
29	300.8	16.7	731	6	AX364381	AX364381 Sequence
30	287	15.9	320427	2	AC108998	AC108998 Rattus no
31	285	15.8	285	6	A78985	A78985 Sequence 4
32	285	15.8	285	6	A91328	A91328 Sequence 4
33	271	15.1	2131	8	AY058869	AY058869 Arabidops
34	260.6	14.5	2170	8	AY060579	AY060579 Arabidops
35	260.4	14.5	2149	8	AY053419	AY053419 Arabidops
36	257.8	14.3	1779	8	AY093961	AY093961 Arabidops
37	227.6	12.6	2391	9	HS081006	U81006 Human p76 m
38	227.6	12.6	2805	6	ARI37623	ARI37623 Sequence
39	227.4	12.6	2068	9	BC021107	BC021107 Homo sapi
40	227.4	12.6	3761	9	BC022850	BC022850 Homo sapi
41	227.4	12.6	4028	9	D87444	D87444 Human mRNA
42	226	12.6	1878	9	AK093507	AK093507 Homo sapi

ALIGNMENTS

RESULT 1  
LOCUS A91337  
DEFINITION Sequence 13 from Patent WO9826065.  
ACCESSION A91337  
VERSION A91337.1 GI:6740338  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1800)  
AUTHORS Lenzen.G. and Strosberg.A.D.  
TITLE MAMMALIAN ICYP (IODOCYANOPINDOLOL) RECEPTOR AND ITS APPLICATIONS  
JOURNAL Patent: WO 9826065-A 13 18-JUN-1998;  
VETIGEN (FR); LENZEN GERLINDE (FR)

linear PAT 22-JAN-2000

FEATURES  
Source Location/Qualifiers  
1..1800  
/organism="unidentified"  
/db\_xref="taxon:32644"  
CDS  
3..1733  
/note="unnamed protein product"  
/codon\_start=1  
/db\_xref="GI:6740339"  
/translation="AALWLLLLLLLPRTRADEHEHTYODKBEVVLWMTVGPYHNRQET  
IKYFLSPCVGSKSISHYHETLGEALQGELEFSGLDIKFKDDMPATGICEIDLDKE  
KDAFVAIKNHVQWYIDDLPIWINGEADENGEDYLMWTKKLEIFGNRIVDV  
NLTSEGVKLVNFKIQMSYVKWKSDKDFEDFKYLDSPFQKHRIHMFISNSEM  
MVLVLGLVLMRTLKDVARYSKEEMDDMDRLDGYGKQVHGDFRPPSSHPL  
IFSLIGSGCOIFAVSLIIVAMLEDLYTERGSMLETAIFVYAAATSPVANGPFGSLY  
ARQGRRIKQMFICAGFLIPANVCGTAFFINFIALIYHASRAIPFGTRWAVCCICFFV  
ILPLNVGTILGRNLSGPNPPCRVNAVPRPIPEKKWFMPEAVIVLGLGILPFGSIFI  
EMFIPTSFMAKYIYVYVPMMLVLILCIVTCVTIVCTVFLNAGEDYRWQMTSFLS  
AASATIIYVMYSFYFFFKMKYGLFQTSFYFGYMAVFESTALGIMCAIGMYGTSAFV  
RKIVTNKVID"  
BASE COUNT 509 a 300 c 400 g 591 t  
ORIGIN  
Query Match 100.0%; Score 1800; DB 6; Length 1800;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1800; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 CCGCCGCGCTGGCTGCTGCTGCTGCTGCCCGCGGACCCGGCGGACGACGAC 60  
DB 1 CCGCCGCGCTGGCTGCTGCTGCTGCTGCCCGCGGACCCGGCGGACGACGAC 60  
QY 61 ACACGTATCAAGATAAAGAGGAAGTGTCTTATGGATGAATACTCTTGGGCCCTACCATA 120  
DB 61 ACACGTATCAAGATAAAGAGGAAGTGTCTTATGGATGAATACTCTTGGGCCCTACCATA 120  
QY 121 ATCTGCAAGAACATATAGTACTTTTCCATTTCTGTGTGGGGTCAAAAAAAGTA 180  
DB 121 ATCTGCAAGAACATATAGTACTTTTCCATTTCTGTGTGGGGTCAAAAAAAGTA 180  
QY 181 TCAGTCATTACCATCAAACTCTGGGAGAGCACTTCAAGGGTTCGAATGGAAATTAGT 240  
DB 181 TCAGTCATTACCATCAAACTCTGGGAGAGCACTTCAAGGGTTCGAATGGAAATTAGT 240  
QY 241 GTCTGGATATTAATTTAAAGATGATGTGATGCCAGCCACTTACGTGCAAAATGATTAG 300  
DB 241 GTCTGGATATTAATTTAAAGATGATGTGATGCCAGCCACTTACGTGCAAAATGATTAG 300  
QY 301 ATAAAGAAAGAGAGATGATTTGTATATGCCATAAAAAATCATTTACGTACCAGATGT 360  
DB 301 ATAAAGAAAGAGAGATGATTTGTATATGCCATAAAAAATCATTTACGTACCAGATGT 360  
QY 361 ACATAGATGATTTACCAATATGGGTATTTGTGGTGGAGCTGATGAAATGGAGAAGATT 420  
DB 361 ACATAGATGATTTACCAATATGGGTATTTGTGGTGGAGCTGATGAAATGGAGAAGATT 420  
QY 421 ACTATCTTTGGACCTATAAAAAACHTTGAATAGGTTTAAATGGAATCGAATTTGTATG 480  
DB 421 ACTATCTTTGGACCTATAAAAAACHTTGAATAGGTTTAAATGGAATCGAATTTGTATG 480  
QY 481 TTAATCTAATAGTGAAGAAAGGTGAACCTGGTTTCCAAATACTAAATCCAGATGTGAT 540  
DB 481 TTAATCTAATAGTGAAGAAAGGTGAACCTGGTTTCCAAATACTAAATCCAGATGTGAT 540  
QY 541 ATTCAGTAAATGGAAGAAAGTCAGATGTGAATTTGAAGATCGATTTGCACAAATATCTTG 600  
DB 541 ATTCAGTAAATGGAAGAAAGTCAGATGTGAATTTGAAGATCGATTTGCACAAATATCTTG 600  
QY 601 ATCCGTCCTTTTCAACATCGGATTCATTTGTTTCAATTTTCAACCTCTTCATGATGG 660  
DB 601 ATCCGTCCTTTTCAACATCGGATTCATTTGTTTCAATTTTCAACCTCTTCATGATGG 660  
QY 661 TGATCTCTTGGTGGGCTTAGTTTCAATGATTTTAAATGAGAACATTAAGAAAGATTATG 720  
DB 661 TGATCTCTTGGTGGGCTTAGTTTCAATGATTTTAAATGAGAACATTAAGAAAGATTATG 720

DB 661 TCATCTCTTGGTGGGCTTAGTTTCAATGATTTTAAATGAGAACATTAAGAAAGATTATG 720  
QY 721 CTCGGTACAGTAAAGAGCAAGAAATGGATGATATGGATAGAGACCTAGGAGATGAATATG 780  
DB 721 CTCGGTACAGTAAAGAGCAAGAAATGGATGATATGGATAGAGACCTAGGAGATGAATATG 780  
QY 781 GATGGAACACAGTGCATGGAGATGTATTTAGACCATCAAGTCAACCACCTGATATTTCTCT 840  
DB 781 GATGGAACACAGTGCATGGAGATGTATTTAGACCATCAAGTCAACCACCTGATATTTCTCT 840  
QY 841 CTCGTATTTGGTTCGGATGTCAGATATTTGCTGTGCTCTCATCGTATATTTGTTGCA 900  
DB 841 CTCGTATTTGGTTCGGATGTCAGATATTTGCTGTGCTCTCATCGTATATTTGTTGCA 900  
QY 901 TCATAGAGATTTATATACTCAGAGGGGATCAATCTCAGTACAGCCATATTTGCTCTATG 960  
DB 901 TCATAGAGATTTATATACTCAGAGGGGATCAATCTCAGTACAGCCATATTTGCTCTATG 960  
QY 961 CTGCTACCTCTCCAGTGAATGTTATTTGGAGGAAGTCTGTATGCTAGACAAGAGGAA 1020  
DB 961 CTGCTACCTCTCCAGTGAATGTTATTTGGAGGAAGTCTGTATGCTAGACAAGAGGAA 1020  
QY 1021 GGAGATGATTAAGCAGATGTTTATTTGGGGCATTCCTTATCCAGCTATGGTGTGCGCA 1080  
DB 1021 GGAGATGATTAAGCAGATGTTTATTTGGGGCATTCCTTATCCAGCTATGGTGTGCGCA 1080  
QY 1081 CTGCCCTCTTCATCAATTTTCATAGCCATTTATACCATGCTTCAAGAGCCATTCCTTTTG 1140  
DB 1081 CTGCCCTCTTCATCAATTTTCATAGCCATTTATACCATGCTTCAAGAGCCATTCCTTTTG 1140  
QY 1141 GAACAATGGTGGCCGTGTTGTCATCTCTTTTGTGTTTATTTCTTCCCTCTAAATCTGTTG 1200  
DB 1141 GAACAATGGTGGCCGTGTTGTCATCTCTTTTGTGTTTATTTCTTCCCTCTAAATCTGTTG 1200  
QY 1201 GTACAATACTTTGGCGGAATCTGTCAGTCAAGCCCAACTTTCTTGTGCTGTCATGCTG 1260  
DB 1201 GTACAATACTTTGGCGGAATCTGTCAGTCAAGCCCAACTTTCTTGTGCTGTCATGCTG 1260  
QY 1261 TGCCTCGTCTATACCGGAGAAAAAATGGTTTCATGGAGCCCTGCGGTATTTGTTGCGCTGG 1320  
DB 1261 TGCCTCGTCTATACCGGAGAAAAAATGGTTTCATGGAGCCCTGCGGTATTTGTTGCGCTGG 1320  
QY 1321 GTGGAATTTTACCTTTTGGTTCATCTTTTATGAAATGATTTTTCATCTTCCAGCTTTTCT 1380  
DB 1321 GTGGAATTTTACCTTTTGGTTCATCTTTTATGAAATGATTTTTCATCTTCCAGCTTTTCT 1380  
QY 1381 GGCATATAAAGATCTATTTATGCTATGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
DB 1381 GGCATATAAAGATCTATTTATGCTATGCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTG 1440  
QY 1441 TTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
DB 1441 TTGTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 GGTGGAATGGAACAAGTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
DB 1501 GGTGGAATGGAACAAGTTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1561 CTTTTTACTACTATTTTTCACAAACAAAGATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
DB 1561 CTTTTTACTACTATTTTTCACAAACAAAGATGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
QY 1621 TTGATATATGCGGTATTTAGCACAGCTTTGGGGATTAATGTGGAGCGATTTGCTTACA 1680  
DB 1621 TTGATATATGCGGTATTTAGCACAGCTTTGGGGATTAATGTGGAGCGATTTGCTTACA 1680  
QY 1681 TGGGAACAAGTGCCTTTGTCCGAAAAATCTACTACTAAATTTGACATAGACGCCA 1740  
DB 1681 TGGGAACAAGTGCCTTTGTCCGAAAAATCTACTACTAAATTTGACATAGACGCCA 1740  
QY 1741 AGAAACCTGGAACCTTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGACACAGAAA 1800  
DB 1741 AGAAACCTGGAACCTTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGACACAGAAA 1800





Db	1095	CTCTGATTGGTTCTCGATGTCAGATATTGTCGTCTCTCTCATCGTTATATATTGTTGCAA	1154
Qy	901	TCATAGAAGATTTATATACTAGAGGGGATCAATGCTCAGTACAGCCATATTGTCATG	960
Db	1155	TCATAGAAGATTTATATACTAGAGGGGATCAATGCTCAGTACAGCCATATTGTCATG	1214
Qy	961	CTGCTACGTCCTCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTAGACAAGGAGAA	1020
Db	1215	CTGCTACGTCCTCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTAGACAAGGAGAA	1274
Qy	1021	GGAGATGATAAAGCAGATGTTTATTGGGCGATTCCTTATCCCAGCTATGGTGTGGCA	1080
Db	1275	GGAGATGATAAAGCAGATGTTTATTGGGCGATTCCTTATCCCAGCTATGGTGTGGCA	1334
Qy	1081	CTGCTCTCTTCATCAATTTTCATAGCCATTTATACCATGCTTCAAGAGCCATTCCTTTG	1140
Db	1335	CTGCTCTCTTCATCAATTTTCATAGCCATTTATACCATGCTTCAAGAGCCATTCCTTTG	1394
Qy	1141	GAACAATGGTGGCGTTTGGTGCATCTGTTTTTGGTATTCTTCCCTCTAAATCTGTGTG	1200
Db	1395	GAACAATGGTGGCGTTTGGTGCATCTGTTTTTGGTATTCTTCCCTCTAAATCTGTGTG	1454
Qy	1201	GTACAATACTTTGGCGGAAATCTGTACGCTCAGCCCAACTTTTCTTGTCTGTCGAATGCTG	1260
Db	1455	GTACAATACTTTGGCGGAAATCTGTACGCTCAGCCCAACTTTTCTTGTCTGTCGAATGCTG	1514
Qy	1261	TGCTCTGCTCTATACCGGAGAAAAATGGTTCATCGAGCCCTGCGGTTATTTGTTGCCCTG	1320
Db	1515	TGCTCTGCTCTATACCGGAGAAAAATGGTTCATCGAGCCCTGCGGTTATTTGTTGCCCTG	1574
Qy	1321	GTGGAATTTTACCTTTTGGTTCATCTTTATTTGAATGATTTTCAATCTCAGCTCTTCT	1380
Db	1575	GTGGAATTTTACCTTTTGGTTCATCTTTATTTGAATGATTTTCAATCTCAGCTCTTCT	1634
Qy	1381	GGGCATATAAGATCTATTATCTCTATGCTTCATGATGCTGGTGTATCTGCTGTACA	1440
Db	1635	GGGCATATAAGATCTATTATCTCTATGCTTCATGATGCTGGTGTATCTGCTGTACA	1694
Qy	1441	TTGTGACTGTCTGTGACTATTGTGTCACATATTTTCTACTAAATGCAGAAGATTACC	1500
Db	1695	TTGTGACTGTCTGTGACTATTGTGTCACATATTTTCTACTAAATGCAGAAGATTACC	1754
Qy	1501	GGTGGCAATGACAAAGTTTCTCTGCTGCATCAACTGCAATCTATGTTTACATGTATT	1560
Db	1755	GGTGGCAATGACAAAGTTTCTCTGCTGCATCAACTGCAATCTATGTTTACATGTATT	1814
Qy	1561	CCTTTTACTACTATTTTTCAAAAAAGATGTATGGCTTATTTCAACATCATTTTACT	1620
Db	1815	CCTTTTACTACTATTTTTCAAAAAAGATGTATGGCTTATTTCAACATCATTTTACT	1874
Qy	1621	TTGGATATATGGCGTATTTAGCAGAGCTTTGGGATAATGTGGAGCGATTGTTTACA	1680
Db	1875	TTGGATATATGGCGTATTTAGCAGAGCTTTGGGATAATGTGGAGCGATTGTTTACA	1934
Qy	1681	TGGGAACAAGTGCCTTTGTCGGAATAATCTATCTAATGTGAAATTTGACTTAGAGACCA	1740
Db	1935	TGGGAACAAGTGCCTTTGTCGGAATAATCTAATGTGAAATTTGACTTAGAGACCA	1994
Qy	1741	AGAAAACCTGGAATTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTGCACAGCAAA	1800
Db	1995	AGAAAACCTGGAATTTGGATCAATTTCTTTTTCATAGGGGTGGAACCTGCACAGCAAA	2054
RESULT 4			
AF160213			
LOCUS	AF160213	3389 bp	mRNA linear PRI 02-MAY-2000
DEFINITION	Homo sapiens endomembrane protein emp70 precursor		
ACCESSION	AF160213		
VERSION	AF160213.1		
KEYWORDS	GI:7677067		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		

[illegible]

```
|||||
Db 474 ACTATCTTTGGACCTATAAAACCTTGAATAGGTTTAAATGGAATCGAATTTGTGTATG 533
QY 481 TTAATCTAACTAGTGAAGGAAAGGTGAAACT--GGTTTCCAAATACATAAAATCCAGATGTCA 539
Db 534 TTAATCTAACTAGTGAAGGAAAGGTGAAACTGGTTTCCAAATACATAATCCAGATGTCA 593
QY 540 TATTCAGTAAATGGAAGGACAGATGTGAATTTTGAAGATCGATTTTGACAAATATCTTT 599
Db 594 TATTCAGTAAATGG--AAAAGTCAGATGTGAATTTTGAAGATCGATTTTGAC--AATATCTTT 651
QY 600 GATCCGCTCCTTTTCAACATCGGATTCATTCGTTTTCATTTTCAATTTTCAACTCCTTCATGATG 659
Db 652 GATC--GTCCCTTTTTCACATCGGATTCATTCGTTTTCATTTTCAATTTTCAACTCCTTCATGATG 709
QY 660 GTGATCTTCTCGTGGCCTTAGTCAATGATTTAATGAGAACATTAAGAAAAGATATAT 719
Db 710 GTGATCTTCTCGTGGCCTTAGTCAATGATTTAATGAGAACATTAAGAAAAGATATAT 769
QY 720 GCTCGGTACAGTAAAGGAGGAAGAAATGGATGATATGGATAGAGACCTAGGAGATGAATAT 779
Db 770 GCTCGGTACAGTAAAGGAGGAAGAAATGGATGATATGGATAGAGACCTAGGAGATGAATAT 829
QY 780 GGATGAAACAGGTGCGATGGAGATGATTTTACACATCAAGTCACCCACTGATATTTTCC 839
Db 830 GGATGAAACAGGTGCGATGGAGATGATTTTACACATCAAGTCACCCACTGATATTTTCC 889
QY 840 TCTCTGATTTGTTCTGAGTCCAGATATTTGCTGTCTCTCTCATCGTATTTATTTCTTCCA 899
Db 890 TCTCTGATTTGTTCTGAGTCCAGATATTTGCTGTCTCTCTCATCGTATTTATTTCTTCCA 949
QY 900 ATGATAGAAGATTTATATACTGAGAGGGGATCAATGCTCAGTACAGCCATATTTGCTCTAT 959
Db 950 ATGATAGAAGATTTATATACTGAGAGGGGATCAATGCTCAGTACAGCCATATTTGCTCTAT 1009
QY 960 GCTGCTACGTCCTCAGTGAATGGTATTTTGGAGGAAGTCTGCTATGCTAGACAAGGAGA 1019
Db 1010 GCTGCTACGTCCTCAGTGAATGGTATTTTGGAGGAAGTCTGCTATGCTAGACAAGGAGA 1069
QY 1020 AGGAGATGATAAAGCAGATGTTTATTTGGGGCATTCCTTATCCCAAGCTATGTGTGTGCG 1079
Db 1070 AGGAGATGATAAAGCAGATGTTTATTTGGGGCATTCCTTATCCCAAGCTATG--GGTGTGC 1127
QY 1080 ACTGCTCTTCTTCATCAATTTATAGCCATTTATTTACCATGCTTCAAGAGCCATTCCTTTT 1139
Db 1128 ACTGCTCTTCTTCATCAATTTATAGCCATTTATTTACCATGCTTCAAGAGCCATTCCTTTT 1187
QY 1140 GGAACAATGGTGGCGGTTTGTGCATCTGTTTTTGTATTTCTTCTCTAAATCTTGTT 1199
Db 1188 GGAACAATGGTGGCGGTTTGTGCATCTGTTTTTGTATTTCTTCTCTAAATCTTGTT 1247
QY 1200 GGTACAATACTTTGGCGGAAATCTGTGAGTCAAGCCAACTTTTCTTGTGCTCAATGCT 1259
Db 1248 GGTACAATACTTTGGCGGAAATCTGTGAGTCAAGCCAACTTTTCTTGTGCTCAATGCT 1307
QY 1260 GTGCTCGCTCTATACCGGAGAAAAAATGGTTTCATGGAGCCCTGCGGTTATTTTGGCCTG 1319
Db 1308 GTGCTCGCTCTATACCGGAGAAAAAATGGTTTCATGGAG--CTGCGGTTATTTTGGCCTG 1366
QY 1320 GGTGGAATTTTACCTTTTGGTTCAATCTTTATTTGAATGATTTTCATCTTACGCTCTTTC 1379
Db 1367 GGTGGAATTTTACCTTTTGGTTCAATCTTTATTTGAATGATTTTCATCTTACGCTCTTTC 1426
QY 1380 TGGGATATAAGATCTATTATGCTCTATGGCTTCATGATGCTGCTGCTGTTATCTCTGTC 1439
Db 1427 TGGGATATAAGATCTATTATGCTCTATGGCTTCATGATGCTGCTGCTGTTATCTCTGTC 1486
QY 1440 ATTGTGACTGCTGTGCTGACTATTTGTGTCACATATTTTCTACTAAATCCAGAGATTAC 1499
Db 1487 ATTGTGACTGCTGTGCTGACTATTTGTGTCACATATTTTCTACTAAATCCAGAGATTAC 1546
QY 1500 CCGTGGCAATGACAAGTTTCTCTGCTGCTCACTCAATCGAATCTATGTTTACATGATAT 1559
|||||
```

```
Db 1547 CGGTGGCAATGACAAGTTTCTCTCTCTGTCATCAACTGCAATCTATGTTTACATGATAT 1606
QY 1560 TCCTTTTACTACTATTTTTCAAAAAAGAGATGATGGCTTATTTTCAAAACATCATTTTAC 1619
Db 1607 TCCTTTTACTACTATTTTTCAAAAAAGAGATGATGGCTTATTTTCAAAACATCATTTTAC 1666
QY 1620 TTTGGATATATGGCGGTATTTTAGCACAGCCCTTTGGGGATAAATGTTGTGGAGCGATTGGTTAC 1679
Db 1667 TTTTGGATATATGGCGGTATTTTAGCACAGCCCTTTGGGGATAAATGTTGTGGAGCGATTGGTTAC 1726
QY 1680 ATGGGAAACAGTGCCTTTGTCGAAAAAATCTATCTAATGTAATAATTCACCTAGAGACCC 1739
Db 1727 ATGGGAAACAGTGCCTTTGTCGAAAAAATCTATCTAATGTAATAATTCACCTAGAGACCC 1786
QY 1740 AAGAAAACTGGAACCTTTGGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGACACAGCAA 1799
Db 1787 AAGAAAACTGGAACCTTTGGGATCAATTTCTTTTTCATAGGGGTGGAACCTTGACACAGCAA 1846
QY 1800 A 1800
Db 1847 A 1847

RESULT 5
AK027788
LOCUS
DEFINITION Homo sapiens cDNA FLJ14882 fis, clone PLACE1003537, weakly similar
to ENDSOMAL P24A PROTEIN PRECURSOR.
ACCESSION AK027788
VERSION AK027788.1 GI:14042725
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens placenta cDNA to mRNA, clone_lib:PLACE1
clone:PLACE1003537.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y.,
Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H.,
Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K.,
Arita,M., Nabekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J.,
Wakamatsu,A., Nakamura,Y., Nagahari,K., Masuho,Y. and Oshima,A.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2459)
Isogai,T. and Otsuki,T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology; cDNA library construction,
5'- & 3'- end one pass sequencing and clone selection; Helix
Research Institute (supported by Japan Key Technology Center etc.)
and Department of Virology, Institute of Medical Science,
University of Tokyo.
FEATURES
Location/Qualifiers
1..2459
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="PLACE1003537"
/tissue_type="placenta"
/clone_lib="PLACE1"
/notes="cloning vector: pME18SFL3"
6..1643
/notes="unnamed protein product"
/codon_start=1
/protein_id="BAB55369.1"
/db_xref="GI:14042726"
/translation="MNTVGPVHNRQETKYVFSLPFCVSKKSIHSHYHETLGEALQGYE
LEPSGLDIKPKDDVMPATYCEIDLDEKRAFVYAIKNHYQMY IDPLPLWGLVGEA
DENGEDYILWTYKLEIFGNRIVDNLTSECKVLVPNTKIQMSYSVKWKSDVRF
```

EDRFKYLDPSEFFQHRHMFISFNSPMWIFVLGLVSMILMRTLRKDYARYSKEEMD  
DMRDLDGDEYKQVHGDFRPSHPLIFSSGLSGCQIFAVSLIVIIIVAMIEDLYTE  
RGSMLSTAIPIVVAATSPVNGFYGSLYAROGGRWIKOMFIGAFLIPAMVCGTFAFPI  
FIATYHASRAIPFGTMVAVCCICFVILPLNLVYGTILGNLSQGNFPCRVNAVPRP  
IPBKNMPEAVLIVLUGLPLPFGSIFIEIMFIFTSEWAKIYIYVGFMMILVLICIV  
TVCVITVCYFLNABEDYRWQWTSFLSAASATAYVMYSFYFYFFFKMYGLPQTSFY  
FGYMAVFSTALGIMCAIGMGTSAFVRKIYTNVKID"  
BASE COUNT 728 a 399 c 490 g 842 t  
ORIGIN

Query Match 94.9%; Score 1708.4; DB 9; Length 2459;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1709; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	91	TATGGATGAATACCTGTGGGCCCTACCATAATCGTCAAGAAACATATAAGTACTTTTCAC	150
DB	1	TATGGATGAATACCTGTGGGCCCTACCATAATCGTCAAGAAACATATAAGTACTTTTCAC	60
QY	151	TTCCATTCTCTGTGGGTCACAAAAAGATATCAGTCAATTACCATGAACCTCTGGGAGAAG	210
DB	61	TTCCATTCTCTGTGGGTCACAAAAAGATATCAGTCAATTACCATGAACCTCTGGGAGAAG	120
QY	211	CACCTCAAGGGTGAATTTAGTGGTCTGGATATAAATTTAAAGATGATGTGA	270
DB	121	CACCTCAAGGGTGAATTTAGTGGTCTGGATATAAATTTAAAGATGATGTGA	180
QY	271	TGCCAGCCACTTACTGTGAATTCATTAGATAAAGAAAGAGAGATGCGATTTGTATATG	330
DB	181	TGCCAGCCACTTACTGTGAATTCATTAGATAAAGAAAGAGAGATGCGATTTGTATATG	240
QY	331	CCATAAAAAATCAATTACTGGTACCAGATGACATAGATGATTTACCAATATGGGGTATTG	390
DB	241	CCATAAAAAATCAATTACTGGTACCAGATGACATAGATGATTTACCAATATGGGGTATTG	300
QY	391	TTGGTGAGGCTGATGAAANTGGAGAACTTACTATCTTTGGACCTATAAAAACCTTGAAA	450
DB	301	TTGGTGAGGCTGATGAAANTGGAGAACTTACTATCTTTGGACCTATAAAAACCTTGAAA	360
QY	451	TAGGTTTAAATGAAATCGAATCTTCATGTTAATCTAATCTAGTGAAGGAAAGGTGAAC	510
DB	361	TAGGTTTAAATGAAATCGAATCTTCATGTTAATCTAATCTAGTGAAGGAAAGGTGAAC	420
QY	511	TGGTTCCAAATACTAAATCCAGATGTCATATTCAGTAAAAATGAAAAAGTCAGATGTGA	570
DB	421	TGGTTCCAAATACTAAATCCAGATGTCATATTCAGTAAAAATGAAAAAGTCAGATGTGA	480
QY	571	AATTTGAAGATCGATTTGACAAATATCTTGATCGGTCCTTTTTCACATCGGATTCATT	630
DB	481	AATTTGAAGATCGATTTGACAAATATCTTGATCGGTCCTTTTTCACATCGGATTCATT	540
QY	631	GGTTTTCAATTTTCAACTCTTCATGATGTTGATCTTCTGGTGGCTTAGTTTCAATGA	690
DB	541	GGTTTTCAATTTTCAACTCTTCATGATGTTGATCTTCTGGTGGCTTAGTTTCAATGA	600
QY	691	TTTTTAATGAGAACATTAAGAAAAGATTATGCTCGGTACAGTAAAGAGAAAGAAATGGATG	750
DB	601	TTTTTAATGAGAACATTAAGAAAAGATTATGCTCGGTACAGTAAAGAGAAAGAAATGGATG	660
QY	751	ATATGGATAGACCTAGGAGATGAATATGGATGGAACACAGGTCATGGAGATGATTTTA	810
DB	661	ATATGGATAGACCTAGGAGATGAATATGGATGGAACACAGGTCATGGAGATGATTTTA	720
QY	811	GACCATCAAGTCACCCACTGATATTTTCCCTCTGATTTGGTCTCGGATGTCAGATATTG	870
DB	721	GACCATCAAGTCACCCACTGATATTTTCCCTCTGATTTGGTCTCGGATGTCAGATATTG	780
QY	871	CTGTGCTCTCATCGTTATTATTGTTGCAATAGTAGAAGATTTATATCTAGCAGGGGAT	930
DB	781	CTGTGCTCTCATCGTTATTATTGTTGCAATAGTAGAAGATTTATATCTAGCAGGGGAT	840
QY	931	CAATGCTCAGTACAGCCATATTTGCTGATGCTGCTAGCTCCAGTGAATGGTTATTTG	990
DB	841	CAATGCTCAGTACAGCCATATTTGCTGATGCTGCTAGCTCCAGTGAATGGTTATTTG	900

QY	991	GAGGAAGCTGTATGCTAGACAAGGAGGAGAGATGATAAAGCAGATGTTTATTGGG	1050
DB	901	GAGGAAGCTGTATGCTAGACAAGGAGGAGAGATGATAAAGCAGATGTTTATTGGG	960
QY	1051	CATTCCCTATCCAGCTATGCTGTGGTACAACTTTGGCCGAAATCTGTTCAGTCT	1110
DB	961	CATTCCCTATCCAGCTATGCTGTGGTACAACTTTGGCCGAAATCTGTTCAGTCT	1020
QY	1111	ATTACCATGCTTCAAGAGCCATTCCCTTTTGAACAATGGTGGCGTTTGTTCATCTGTT	1170
DB	1021	ATTACCATGCTTCAAGAGCCATTCCCTTTTGAACAATGGTGGCGTTTGTTCATCTGTT	1080
QY	1171	TTTTTTGTTATTTCCCTCTAAATCTTGTGTGTAACATTTGGCCGAAATCTGTTCAGTCT	1230
DB	1081	TTTTTTGTTATTTCCCTCTAAATCTTGTGTGTAACATTTGGCCGAAATCTGTTCAGTCT	1140
QY	1231	AGCCAACTTTCCCTCTGCTCAATGCTGTGCTCTCTATACCCGAGAGAAAAATGGT	1290
DB	1141	AGCCAACTTTCCCTCTGCTCAATGCTGTGCTCTCTATACCCGAGAGAAAAATGGT	1200
QY	1291	TCATGGAGCCTGGGTTTATTTGTTGCTGGTGGTGAATTTTACCTTTTGGTTCATCTTTA	1350
DB	1201	TCATGGAGCCTGGGTTTATTTGTTGCTGGTGGTGAATTTTACCTTTTGGTTCATCTTTA	1260
QY	1351	TTGAAATGTTATTTCACTTTCACGCTTTTCTGGGCATATAAGATCTATTATGCTATGGCT	1410
DB	1261	TTGAAATGTTATTTCACTTTCACGCTTTTCTGGGCATATAAGATCTATTATGCTATGGCT	1320
QY	1411	TCATGATGCTGGTGGTGTATCTGTCATTTGTGACTGTGTCGTGACATTTGTCGTGCA	1470
DB	1321	TCATGATGCTGGTGGTGTATCTGTCATTTGTGACTGTGTCGTGACATTTGTCGTGCA	1380
QY	1471	CATATTTTCTACTAAATGCGAGAAATTTACCGTGGCAATGGACAAAGTTTCTCTCTGCTG	1530
DB	1381	CATATTTTCTACTAAATGCGAGAAATTTACCGTGGCAATGGACAAAGTTTCTCTCTGCTG	1440
QY	1531	CATCAACTGCAATCTATGTTTACATGATTTCTTTTACTACTATTATTTTCAAAACAAGA	1590
DB	1441	CATCAACTGCAATCTATGTTTACATGATTTCTTTTACTACTATTATTTTCAAAACAAGA	1500
QY	1591	TGATGCTCTATTTCAAACTATCTTTTACTTTTGGATATATGGCGGTATTTAGCACAGCCT	1650
DB	1501	TGATGCTCTATTTCAAACTATCTTTTACTTTTGGATATATGGCGGTATTTAGCACAGCCT	1560
QY	1651	TGGGTAATGCTGTGGAGCGATTTGGTTACATGGGAAACAGTGCCTTTGTCGGAATAATCT	1710
DB	1561	TGGGTAATGCTGTGGAGCGATTTGGTTACATGGGAAACAGTGCCTTTGTCGGAATAATCT	1620
QY	1711	ATACTAATGTAATAATTTGACTAGAGCCCAAGAAACCTGGAACCTTTGGATCAATTTCTT	1770
DB	1621	ATACTAATGTAATAATTTGACTAGAGCCCAAGAAACCTGGAACCTTTGGATCAATTTCTT	1680
QY	1771	TTTCATAGGGGTGGAACTTGCACAGCAAAA	1800
DB	1681	TTTCATAGGGGTGGAACTTGCACAGCAAAA	1710

RESULT 6  
BC020959  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

BC020959  
Homo sapiens, clone MGC:8842 IMAGE:3850015, mRNA, complete cds.  
BC020959  
BC020959.1 GI:18088072  
MGC.  
Homo sapiens.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 2870)  
Strausberg,R.  
Direct Submission  
Submitted (03-JAN-2002) National Institutes of Health, Mammalian



Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgapps@mail.nih.gov](mailto:cgapps@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Center, Stanford University School of Medicine, Stanford, CA 94305  
Web site: <http://www.shgc.stanford.edu>  
Contact: (Dickson, Mark) [mcdepaxil.stanford.edu](mailto:mcdepaxil.stanford.edu)  
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAK plate: 20 Row: 1 Column: 13.  
Location/Qualifiers

FEATURES

source  
1..2870  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="MGC:8842 IMAGE:3850015"  
/tissue\_type="Colon, adenocarcinoma"  
/clone\_lib="NIH\_MGC\_65"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6"  
191..1567  
/codon\_start=1  
/product="Unknown (protein for MGC:8842)"  
/protein\_id="AAH20959.1"  
/db\_xref="GI:18088073"  
/translation="MYIDPLPIWIGVEADENGEDYLYWTKKLEIGNRIVDVNL  
TSEGVKLPVNTKIQMSYVKWKSDFVDFRDKYLFDSFFQRIHWFSPNSMMV  
IFLVLGSLMLRTLRKYDIARYSKSEEDMDRLDGLDEYGNKQVHGDVFRPSHPLIF  
SSLIGSGQIFAVSLIVIAVIEDLYTERGSMLETAIFVVAATSPVNGYGGSLYAR  
QGGRWIKQMFGLIFPAMVCGTAFFINFIAIYHASKRAIPFGFMVAVCCICFFVIL  
PLNLVGTILGRNLGQNPFCPCRVNAPRPIPKWFMPEPAVIVCLGGLPGSIFIEI  
YEIFTSWAYKLYVYGPMVLVLILCIIVTCVTIVCTYFELLNAEDYRQWTSFLSAA  
STAIYVWHSFYFFFKYMKVGLPQTSFYFGYMAVFSALGIMCAIGMGTSAFVRK  
IYTNKKID"

CDS

BASE COUNT 866 a 447 c 568 g 989 t

Query Match 90.7%; Score 1632.4; DB 9; Length 2870;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1633; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 167 GTCAAAAAAAGTATCAGTCATTACCATGAAACTCTGGGAGAGACACTTCAAGGGGTGA 226  
Db 1 GTCAAAAAAAGTATCAGTCATTACCATGAAACTCTGGGAGAGACACTTCAAGGGGTGA 60  
QY 227 ATTGGAAATTTAGTGGTCTGGATATTAATTTAAAGATGATGATGCCAGCCACTTACTG 286  
Db 61 ATTGGAAATTTAGTGGTCTGGATATTAATTTAAAGATGATGATGCCAGCCACTTACTG 120  
QY 287 TGAATTTGATTTAGATAAAGAAAGAGATGCATTTGTATATGCCATAAAAAATCATTA 346  
Db 121 TGAATTTGATTTAGATAAAGAAAGAGATGCATTTGTATATGCCATAAAAAATCATTA 180  
QY 347 CTGGTACCAGATGTACATAGATGATTTACCAATATGGGTATTTGTTGGTACGGCTGATCA 406  
Db 181 CTGGTACCAGATGTACATAGATGATTTACCAATATGGGTATTTGTTGGTACGGCTGATCA 240  
QY 407 AAATGAGAAGATTACTATCTTTGGACCTATATAAAACCTTGAATAGGTTTAAATGGAAA 466  
Db 241 AAATGAGAAGATTACTATCTTTGGACCTATATAAAACCTTGAATAGGTTTAAATGGAAA 300  
QY 467 TCGAATTTGTTGATGTTAACTTAACTAGTGAAGGAGGTAAGCTGGTCCCAAACTACTAA 526  
Db 301 TCGAATTTGTTGATGTTAACTTAACTAGTGAAGGAGGTAAGCTGGTCCCAAACTACTAA 360

QY 527 AATCCAGATGTCATATTCAGTAAAAATGAAAAAGTCAGATGTGAAATTTGAAGATCGATT 586  
Db 361 AATCCAGATGTCATATTCAGTAAAAATGAAAAAGTCAGATGTGAAATTTGAAGATCGATT 420  
QY 587 TCACAAATATCTTGATCCGTCCTTTTTCACACATCGGATTCATGGTTTTCATTTTCAA 646  
Db 421 TCACAAATATCTTGATCCGTCCTTTTTCACACATCGGATTCATGGTTTTCATTTTCAA 480  
QY 647 CTTCTTCATGATGGTGATCTCTTGGTGGCTTAGTTTCAATGATTTTAAATGAGAATTT 706  
Db 481 CTTCTTCATGATGGTGATCTCTTGGTGGCTTAGTTTCAATGATTTTAAATGAGAATTT 540  
QY 707 AAGAAAGATTTATGCTCGGTACAGTAAAGAGGAAGAAATGATGATATGATAGAGACCT 766  
Db 541 AAGAAAGATTTATGCTCGGTACAGTAAAGAGGAAGAAATGATGATATGATAGAGACCT 600  
QY 767 AGAGATGAATATGGATGGAACAGGTGCATGGAGATGTATTTAGACCATCAAGTCACCC 826  
Db 601 AGAGATGAATATGGATGGAACAGGTGCATGGAGATGTATTTAGACCATCAAGTCACCC 660  
QY 827 ACTGATATTTTCTCTCTGATGGTTCTGATGTCAGATATTTGCTGCTCTCTCATCTG 886  
Db 661 ACTGATATTTTCTCTCTGATGGTTCTGATGTCAGATATTTGCTGCTCTCTCATCTG 720  
QY 887 TATTATTTGCAATGATAGAAATTTATATCTAGAGGGGATCAATGCTCAGTACACG 946  
Db 721 TATTATTTGCAATGATAGAAATTTATATCTAGAGGGGATCAATGCTCAGTACACG 780  
QY 947 CATATTTGCTATGCTGCTACGTCCTCCAGTGAATGGTTATTTTGGAGGAGTCTGTATGC 1006  
Db 781 CATATTTGCTATGCTGCTACGTCCTCCAGTGAATGGTTATTTTGGAGGAGTCTGTATGC 840  
QY 1007 TAGACAAGGAGGAAGAGATGGATAAAGCAGATGTTTATTTGGGGCATTCCTTATCCACG 1066  
Db 841 TAGACAAGGAGGAAGAGATGGATAAAGCAGATGTTTATTTGGGGCATTCCTTATCCACG 900  
QY 1067 TATGGTGTGTCGACACGCTCTTCTCATCAATTTTCATAGCCATTTATACCATGCTTCAAG 1126  
Db 901 TATGGTGTGTCGACACGCTCTTCTCATCAATTTTCATAGCCATTTATACCATGCTTCAAG 960  
QY 1127 AGCCATTCCTTTTGGAAACAATGGTGGCCCGTTTGTGTCATCTGTTTTTGTATTCTTCC 1186  
Db 961 AGCCATTCCTTTTGGAAACAATGGTGGCCCGTTTGTGTCATCTGTTTTTGTATTCTTCC 1020  
QY 1187 TCTAAATCTGTTGTGTAACAATCTTGGCCGAAATCTGTCAGGTCAGCCCAACTTCTCTTG 1246  
Db 1021 TCTAAATCTGTTGTGTAACAATCTTGGCCGAAATCTGTCAGGTCAGCCCAACTTCTCTTG 1080  
QY 1247 TCGTGCAATGCTGCGCTCGCTCTATACCGGAGAAAAAATGGTTTCATGGAGCCTCGGT 1306  
Db 1081 TCGTGCAATGCTGCGCTCGCTCTATACCGGAGAAAAAATGGTTTCATGGAGCCTCGGT 1140  
QY 1307 TATTGTTTCCCTGGGTGGAAATTTTACCTTTTGGTTCAAATCTTTTAAATGATATTTCAT 1366  
Db 1141 TATTGTTTCCCTGGGTGGAAATTTTACCTTTTGGTTCAAATCTTTTAAATGATATTTCAT 1200  
QY 1367 CTTTCAGTCTTCTTGGGCATATAAAGATCTATTATGTCATGGCTTCATGATGCTGTGCT 1426  
Db 1201 CTTTCAGTCTTCTTGGGCATATAAAGATCTATTATGTCATGGCTTCATGATGCTGTGCT 1260  
QY 1427 GGTTATCTCTGTCATGTCGACTGTCGTGACTATTTGTGTCACATATTTTCTACTAAA 1486  
Db 1261 GGTTATCTCTGTCATGTCGACTGTCGTGACTATTTGTGTCACATATTTTCTACTAAA 1320  
QY 1487 TCGAAGATTTACCGGTGGCAATGGCAAGTTCCTCTCTGCTGTCATCAACTGCAATCTA 1546  
Db 1321 TCGAAGATTTACCGGTGGCAATGGCAAGTTCCTCTCTGCTGTCATCAACTGCAATCTA 1380  
QY 1547 TGTTCATGATTTTCTCTTACTACTATTTTCAAAACAAAGATGATGGCTTATTTTCA 1606  
Db 1381 TGTTCATGATTTTCTCTTACTACTATTTTCAAAACAAAGATGATGGCTTATTTTCA 1440  
QY 1607 AACATCATTTTACTTTTGGATATATGGCGGTATTTAGCACAGCCTTGGGGATTAATGTGTGG 1666



```

|||||
Db 1441 AACATCATTTTACTTGGATATATGGCGTATTTAGCACAGCCTTGGGATAAATGCTGGG 1500
|||||
Qy 1667 AGCGATTGCTTACATGGCAACAAGTCTTTGTCCGAAAAATCTATACTAATGTGAATA 1726
|||||
Db 1501 AGCGATTGCTTACATGGCAACAAGTCTTTGTCCGAAAAATCTATACTAATGTGAATA 1560
|||||
Qy 1727 TGACTAGAGACCCCAAGAAACCTTGGAACTTTGGATCAATTTCTTTTTCATAGGGGTGGAA 1786
|||||
Db 1561 TGACTAGAGACCCCAAGAAACCTTGGAACTTTGGATCAATTTCTTTTTCATAGGGGTGGAA 1620
|||||
Qy 1787 CTTGCACAGCAAAA 1800
|||||
Db 1621 CTTGCACAGCAAAA 1634

RESULT 7
AF269151 Mus musculus transmembrane protein TM9SF3 (Tm9sf3), mRNA linear ROD 09-AUG-2000
DEFINITION
cds.
ACCESSION AF269151
VERSION AF269151.1 GI:9755052
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3143)
Warner, S.J. and Lomax, M.I.
Evolution of the TM9 Super Family of Membrane Spanning Proteins
Unpublished
2 (bases 1 to 3143)
Warner, S.J. and Lomax, M.I.
Direct Submission
Submitted (10-MAY-2000) Kresge Hearing Research Institute,
University of Michigan, 9301E MSRB III, 1150 West Medical Center
Drive, Ann Arbor, MI 48109-0648, USA
FEATURES
source 1..3143
/organism="Mus musculus"
/db_xref="taxon:10090"
1..3143
/gene="Tm9sf3"
83..1846
/gene="Tm9sf3"
/note="transmembrane protein 9 superfamily member 3"
/codon_start=1
/product="transmembrane protein TM9SF3"
/protein_id="AAF98160.1"
/db_xref="GI:9755053"
/translation="MRPLPGAPGVAAAAALLLLLPRASDEHEHTYODKEEYVLNWN
TVGPYINROETKYFSLPCVGSKKSISHYHETLGLALQGVLEFSLGDIKFKDDVHP
GYCEIDLDERDAFYAIKNHIWYOMKIDLPITWIGVEADENEDGYLWYTKKLE
IGPNRNIVDNITSEGVKLVNPKNTQMSYVKWKSVDKFDKDLDPFFQHR
IHWFSIFNMVIFVLGVSLMTRLRDYARYSKEEEMDDMDRLDGLDEYQWQVH
GDVRFSSHLPIFSSLIGSCQIFAVSLIIVIAMIEDLYTERGSLSTAIIFYAATS
PVNGYFGSLYARQGRRTKOMFICAGFLIPAMVCGTAFFINFIAYLYHASRAIPGT
MVAVCCICFEVLPNLVGTILGRNLNLSGQPNFCRVNAPRPIPEKKWFNEPAVIVCL
GGLPFGSFIEMFIFTSFWAYKIYYVGFMMVLVILVICITVCVTVCYFELLNAE
DYMONTSELSAASAIYVMYSFYFFKTKMYGLFQTSFYFGYNAVFSTALGIMCG
AIGMGTSAFVRKIYNFKVID"
BASE COUNT 868 a 564 c 699 g 1012 t
ORIGIN

Query Match 86.2%; Score 1551; DB 10; Length 3143;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1644; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 1 CCGCCGGCTGGCTGCTGCTGCTGCCCGGACCGCGGGGAGGACGACGACCAAC 60
|||||
Db 114 CGCGCGCCGCGGCTGTTGTTGCTGCTGCCCGCGGCTCGGCTGTCGACGACCAAC 173

```

```

Qy 61 ACACGTATCAAGATAAAGAGGAAGTTGTCTTATGGATGAATACTGTTGGGCCCTACCAT 120
|||||
Db 174 ACAGGTATCAAGATAAAGAGGAAGTTGTTTATGGATGAATACTGTTGGGCCCTACCA 233
|||||
Qy 121 ATCGTCAAGAAACATATATAAGTACTTTTCCACTTCCATTTCTGTGGGGTCAAAAAAGTA 180
|||||
Db 234 ATCGCAAGAAACATATATAATACTTTTCCACTTCCATTTCTGTGGGGTCAAAAAAGTA 293
|||||
Qy 181 TCAGTCATTACCATCAAACTCTGGGAGAGACACTTCAAGGGTTCAAATTTGAATTTAGTG 240
|||||
Db 294 TCAGTCATTACCATCAAACTTTGGGAGAGACACTTCAAGGAGTTGAATTTAGATTTAGTG 353
|||||
Qy 241 GTCTGGATATTAATAATTTAAAGATGATGTGATGCCAGCCACTTACTGTCAAAATTTGATTTAG 300
|||||
Db 354 GCCTGGACATTTAAATTTAAAGATGATGTGATGCCAGCCACTTACTGTCAAAATTTGATTTAG 413
|||||
Qy 301 ATAAAGAAAGAGAGATGCATTTGTATATGCCATAAAAAATCATTTACTTGGTACACAGATGT 360
|||||
Db 414 ATAAAGAGAGAGAGATGCGTTTGTGTATGCTATCAAGAAATCACTATTTGGTATCAGATGT 473
|||||
Qy 361 ACATAGATGATTTACCAATATGGGCTATTGTTGGTGAAGCTGATGAAATGGAGAAGATT 420
|||||
Db 474 ACATAGATGACTTACCAATATGGGCTATTGTTGGTGAAGCAGATGAAATGGGAAGATT 533
|||||
Qy 421 ACTATCTTTGGACCTATATAAAAACTTGAATAGGTTTAAATGGAAATCGAATTTGTTGATG 480
|||||
Db 534 ACTATCTTTGGACCTACAAAAACTTGAATAGGCTTTAAATGGAAATCGAATTTGTTGATG 593
|||||
Qy 481 TTAATCTAACTAGTAGGAAGAAAGGTGAACTGGTGTCCAAATCTATAAAATCCAGATGTCA 540
|||||
Db 594 TTAATCTAACTAGTAGGAAGAAAGGTCAAGCTGGTGTCCAAATCTATAAAATACAGATGTCA 653
|||||
Qy 541 ATTCAGTAAATGGAAGAAAGTCAGATGTGAATTTGAAGATCGATTTGACAAATATCTTG 600
|||||
Db 554 ATTCGGTAAATGGAAGAAATCGGATGTAAAAATTTGAAGATCGGATTCGATTAATATCTTG 713
|||||
Qy 601 ATCCGTCCTTTTCAACATCGGATTCATTTGTTTCAATTTTCAACPCCTTCATGATGG 660
|||||
Db 714 ATCCATCCTTTTTCAGCATAGGATTCATCTGTTTCAATTTTCAATTTTCAATTTTCAATTTG 773
|||||
Qy 561 TGATCTCTTCTGGTGGCTTAGTTTCAATGATTTTAAAGAGAACATTAAGAAAGATTAAG 720
|||||
Db 774 TGATCTCTTCTAGTGGGATTTCAATGATTTTAAAGAGAACATTAAGAAAGATTAAG 833
|||||
Qy 721 CTCGTACAGTAAAGAGAAAGTGAATGATATGGATAGAGACCTAGAGATGAATATG 780
|||||
Db 834 CCCGATACAGTAAAGAGAAAGTGAATGATGACATGGAGAGACCTAGAGACGAGTATG 893
|||||
Qy 781 GATGGAACAGGTGCATGGAGATGTATTTAGACCATCAAGTCACCCACTGATATTTTCCT 840
|||||
Db 894 GCTGGAAGCAGGTGCATGGAGATGTGTTACAGCCGTCGAAGTCACCCCTGATCTTCCT 953
|||||
Qy 841 CTCGTATTGTTCTGGATGTCAGATATTTGCTGTGCTCTCATCGTTATTTATTTGTCGA 900
|||||
Db 954 CCTTCATTGGCTCTGGATGTCAGATATTTGCTGTGCTCTCATTTGTTATTTGTTGCA 1013
|||||
Qy 901 TGATAGAGATTTATATACTCAGAGGGATCAATGCTCAGTACAGCCATATTTGTCATATG 960
|||||
Db 1014 TGATAGAGACTTATATACAGAGGGGTCAATGCTCAGCAGACGACCATATTTGTCATATG 1073
|||||
Qy 961 CTGCTAGCTCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTAGACAAGAGGAA 1020
|||||
Db 1074 CTGCTACATCCCAAGTGAACGGGTACTTTTGGAGGAAGTCTGTATGCTAGACAAGAGGAA 1133
|||||
Qy 1021 GGAGATGGATAAAGCAGATGTTTATTTGGGCAATTCCTTATATCCCAAGCTATGTTGTGGCA 1080
|||||
Db 1134 GGAGATGGATAAAGCAGATGTTTATCGSGGCATTCCTTATCCCAAGCTATGTTGTGGCA 1193
|||||
Qy 1081 CTGCTCTCTCATCAATTTCAATGCCATTTATTAACCATTTTCAAGAGCCATTCCTTTTG 1140
|||||
Db 1194 CTGATTTCTTCATCAACTTTTATAGCCATTTTATACCATTCCTAGAGCCATTCCTTTTG 1253
|||||
Qy 1141 GAACAATGGTGGCGCTTGTGTCATCTGTTTTTTTGTGTTATTTCTTCTCTAAATCTTGTG 1200

```

|||||  
Db 1254 GAACAATGGTGGCGTTGTTGGCATCTGTTTCTTGTGTTATCTCTTAAATCTCGTTG 1313  
Qy 1201 GTACAATACTGGCGGAAATCTGTGAGTCCAGTCCAGCCAACTTTCCTGTGCTCAATCGTG 1260  
Db 1314 GTACAATACTGGGTGGAATCTGTGGGTGAGCCCACTTCCCTGTGCTCAATGCCG 1373  
Qy 1261 TGCCTCGTCTATACCGGAGAAAATAGTTTCATGAGCCCTGCGGTTATTGTTTGCCTGG 1320  
Db 1374 TGCCTCGTCCCATCCCGAGAAAAATGGTTATTAGAGCCTGCACTTATTGTTTGCCTGG 1433  
Qy 1321 GTGGAATTTTACCTTTGGTTCAATCTTATTGAAATGATTTATCATCTTCACGTCCTTCT 1380  
Db 1434 GAGGAATTTTACCTTTGGGATCAATCTTATTGAAATGATTTATCATCTTCACGTCCTTCT 1493  
Qy 1381 GGGCATATAAGATCTATTATGCTCTATGCTTCATGCTTCATGCTGGTCTGTTATCTGTGCA 1440  
Db 1494 GGGCATATAAGATCTATTATGCTCTATGCTTCATGCTGGTCTGTTATCTGTGCA 1553  
Qy 1441 TTGTGACTGTCTGTGACTATTGTGTGCACATATTCTTACTAAATGCAGAGATTACC 1500  
Db 1554 TTGTGACTGTCTGTGACTACCTGTGTGCACATATTCTTCTGCTAAAGCGAGAGATTACA 1613  
Qy 1501 GTGGCAATGGACAAGTTTCTCTCTGCTGCATCACTCACTCAATCTATGTTTACATGATT 1560  
Db 1614 GTGGCAGTGGACAGTTTCTCTGCTGCGGCTGCGCGGATCTACGTTTACATGACT 1673  
Qy 1561 CCTTTACTACTATTTTCAAAACAAGATGATGATGCTTATTCAACATCATTTTACT 1620  
Db 1674 CCTTTACTACTATTTTCAAAACAAGATGATGATGCTTATTCAACATCATTTTACT 1733  
Qy 1621 TTGGATATATGGGGTATTATTAGCACAGCCTTGGGGATAATGTGTGGAGCATTTGGTTACA 1680  
Db 1734 TTGGCTACATGGCGTGTATTAGCACACCTTGGGGATAATGTGTGGAGCAATCGTTACA 1793  
Qy 1681 TGGGAACAAGTGCCTTTGTCGGAATAATCTATATACTAATGTGAAAATTTGACTAGAGACCCA 1740  
Db 1794 TGGGAACAAGTGCCTTTGTCGGAATAATCTATATACTAATGTGAAAATTTGACTAGATATCTG 1853  
Qy 1741 AGAAACCTGGAACCTTTGGGATCAATTTCTTTTCATAGGGTGAACCTTCACAGCAAA 1799  
Db 1854 AGACAACATGGAACCTTCGGATCAGTTTCTCTTCATGAGGTGGAACCTTCACAGCAAAA 1912

RESULT 8  
AK000756  
LOCUS AK000756 2973 bp mRNA linear PRI 22-FEB-2000  
DEFINITION Homo sapiens cDNA FLJ20749 fis, clone HEP05301.  
ACCESSION AK000756  
VERSION AK000756.1 GI:7021041  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens hepatoma cell\_line:HepG2 cDNA to mRNA, clone\_lib:HEP  
clone:HEP05301.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (sites)  
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,  
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,  
Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,  
Nakamura,Y., Isogai,T. and Sugano.S.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2973)  
AUTHORS Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,  
Shibahara,T., Tanaka,T. and Nakamura,Y.  
Direct Submission  
TITLE Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Department of Virology; Shirokane-dai, 4-6-1,  
Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@ims.u-tokyo.ac.jp,  
Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan; cDNA full insert

sequencing: Research Association for Biotechnology; cDNA library  
construction, 5'- & 3'-end one pass sequencing: Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency). Location/Qualifiers  
1. .2973  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HEP05301"  
/cell\_line="HepG2"  
/cell\_type="hepatoma"  
/clone\_lib="HEP"  
/notes="cloning vector pME18SFL3"  
14. .1390  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAA91362.1"  
/db\_xref="GI:7021042"  
/translation="MYIDDLPIWIGVEADENGEDYYLWYKKLETFNGNRIVDNL  
TSEGKVLVPTNKIOMSYSVKKKSDVKFEDREKYLDPFFOHRHWFISFNSPMTV  
IELVGLYSMLMRTLRKYARYSKSEEMDDMDRLDGEYGNKQVHGDVPRPSSHLPIF  
SSLIGSCQIFAVSLIIVIAMIEDLYTERGSLSTAFIYATASVPNGVPGGSLXAR  
QGRRWIKQMFAGFLIPAMVCGTAFFINFIAIYHASRAIPGTVWAVCCGCFVIL  
PUNLVGTILGRNLGQPNFPCRVNAPRPIPEKKWMEPAVIVGCGILPFGSFIEM  
YFIETSEWAKIVYVYGEMMLVLILVTCVTIVCTYFLNABDYRQWTSLSAA  
STAIYVMYSFYFFPKMYGLSQTSPFYGVMAVFSTALGIMCGAIGYMGTSAFVRK  
ITYNVRKD"  
BASE COUNT 869 a 476 c 595 g 1033 t  
ORIGIN  
Query Match 80.6%; Score 1450.6; DB 9; Length 2973;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1453; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 344 TTACTGTACCAGATGTACATAGATGATTTACCAATATGGGTATTGTTGGTGAGCTGA 403  
Db 1 TTACTGTGTACCAGATGTACATAGATGATTTACCAATATGGGTATTGTTGGTGAGCTGA 60  
Qy 404 TGAANAATGGAGAAGATTACTATCTTTGGACCTATAAAAACCTTCAAAATAGGTTTAAATGG 463  
Db 61 TGAANAATGGAGAAGATTACTATCTTTGGACCTATAAAAACCTTCAAAATAGGTTTAAATGG 120  
Qy 464 AAATCGAATTTGTTGATGTTAATCTAAGTGAAGGAAAGTGAAACCTGTTCCAAATAC 523  
Db 121 AAATCGAATTTGTTGATGTTAATCTAAGTGAAGGAAAGTGAAACCTGTTCCAAATAC 180  
Qy 524 TAAATCCAGATGTCATATTTCAGTAAATGGAAGGAAAGTCAGATGTGAAATTTGAAGATCG 583  
Db 181 TAAATCCAGATGTCATATTTCAGTAAATGGAAGGAAAGTCAGATGTGAAATTTGAAGATCG 240  
Qy 584 ATTTGACAAATATCTGATCCGCTCTTTTCAACATCGGATTCATTTGTTTCAATTTT 643  
Db 241 ATTTGACAAATATCTGATCCGCTCTTTTCAACATCGGATTCATTTGTTTCAATTTT 300  
Qy 644 CAACCTCTTCATGATGTTGATCTTTGTTGGGCTTCTAGTTCAATGATTTTAAATGAGAAC 703  
Db 301 CAACCTCTTCATGATGTTGATCTTTGTTGGGCTTCTAGTTCAATGATTTTAAATGAGAAC 360  
Qy 704 ATTAAGAAAAGATTATGCTCGGTACAGTAAAGGAGGAAAGAAATGGATGATATGATAGAGA 763  
Db 361 ATTAAGAAAAGATTATGCTCGGTACAGTAAAGGAGGAAAGAAATGGATGATATGATAGAGA 420  
Qy 764 CCTAGAGATGAATATGATGGAAGGAAAGGTCATGAGATGATTTAGACCAATCAATCA 823  
Db 421 CCTAGAGATGAATATGATGGAAGGAAAGGTCATGAGATGATTTAGACCAATCAATCA 480  
Qy 824 CCCACTGATATTTTCCCTCTCTGATTTGGTCTCGGATGTCAGATATTTGCTGTCTCTCAT 883  
Db 481 CCCACTGATATTTTCCCTCTCTGATTTGGTCTCGGATGTCAGATATTTGCTGTCTCTCAT 540  
Qy 884 CGTTATTTGTTGCAATGATAGAACATTTATATCTAGAGGGGATCAATGCTCAGTAC 943  
|||||

FEATURES  
source

CDS

Db	541	CGTTATTATTGTTTCAATAGTAGAAGATTTATATACTAGAGGGCATCAATGCTCAGTAC	600
QY	944	AGCCATATTTTCTATGCTGCTCCTCCAGTGAATGGTTATTATTTGGAGGAAGTCTGTA	1003
Db	601	AGCCATATTTGCTATGCTGCTCCTCCAGTGAATGGTTATTATTTGGAGGAAGTCTGTA	660
QY	1004	TGCTAGACAAGGAGGAGAGATGGATAAAGCAGATGTTTATTTGGGCAATTCCTTATGCC	1063
Db	661	TGCTAGACAAGGAGGAGAGATGGATAAAGCAGATGTTTATTTGGGCAATTCCTTATGCC	720
QY	1064	AGCTATGCTGCTGGCACTGCTTCTTCATCAATTTATACACCATTTATTTACCATGCTTC	1123
Db	721	AGCTATGCTGCTGGCACTGCTTCTTCATCAATTTATACACCATTTATTTACCATGCTTC	780
QY	1124	AAGAGCCATTCCTTTTGGAAACAATGGTGGCGTTTGTGTCATCTGTTTTTTGTTATTTCT	1183
Db	781	AAGAGCCATTCCTTTTGGAAACAATGGTGGCGTTTGTGTCATCTGTTTTTTGTTATTTCT	840
QY	1184	TCCTCTAAATCTGTTTGGTACAATACCTTTGGCCGAAATCTGTCAGGTGAGCCCAACTTTCC	1243
Db	841	TCCTCTAAATCTGTTTGGTACAATACCTTTGGCCGAAATCTGTCAGGTGAGCCCAACTTTCC	900
QY	1244	TTGTCGCTCAATGCTGCTGCTCTATACCGGAGAAAAATGGTTCATGGAGCCCTGC	1303
Db	901	CTGTCGCTCAATGCTGCTGCTCTATACCGGAGAAAAATGGTTCATGGAGCCCTGC	960
QY	1304	GGTTATTGTTTGGCTGGGTGGAATTTTACCTTTTGGTTTCAATCTTTATTTGAAATGTAAT	1363
Db	961	GGTTATTGTTTGGCTGGGTGGAATTTTACCTTTTGGTTTCAATCTTTATTTGAAATGTAAT	1020
QY	1364	CATCTTCAGCTCTTCTTGGGCAATAGATCTATTTATGCTATGCTTCATGGCTTCATGCTGGT	1423
Db	1021	CATCTTCAGCTCTTCTTGGGCAATAGATCTATTTATGCTATGCTTCATGGCTTCATGCTGGT	1080
QY	1424	GCTGGTATCTCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1483
Db	1081	GCTGGTATCTCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1140
QY	1484	AAATGCAAGAAGATTACCGGTGGCAATGGCAAGTTTCTCTCTGCTGCTGCTGCTGCTGCT	1543
Db	1141	AAATGCAAGAAGATTACCGGTGGCAATGGCAAGTTTCTCTCTGCTGCTGCTGCTGCTGCT	1200
QY	1544	CTATGTTTACATGATTTCTTTTACTACTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1603
Db	1201	CTATGTTTACATGATTTCTTTTACTACTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1260
QY	1604	TCAACATCATTTTACTTTTGGATATAGCGGTATTTAGCACACCTTTGGGATAATGTG	1663
Db	1261	TCAACATCATTTTACTTTTGGATATAGCGGTATTTAGCACACCTTTGGGATAATGTG	1320
QY	1664	TGGAGCATTTGGTTACATGGGAACAAGTGGCTTTGTCCGAAAAATCTATATAATGTGAA	1723
Db	1321	TGGAGCATTTGGTTACATGGGAACAAGTGGCTTTGTCCGAAAAATCTATATAATGTGAA	1380
QY	1724	AATGACTAGACCCCAAGAAACCTGGAACTTTGGATCAATTTCTTTTTCATAGGGGTG	1783
Db	1381	AATGACTAGACCCCAAGAAACCTGGAACTTTGGATCAATTTCTTTTTCATAGGGGTG	1440
QY	1784	GAATTCACACGACAAA 1800	
Db	1441	GAATTCACACGACAAA 1457	
RESULT	9		
BC004799			
LOCUS	BC004799	2848 bp	mRNA linear ROD 07-AUG-2002
DEFINITION	Mus musculus, clone MGC:8136 IMAGE:3589474, mRNA, complete cds.		
ACCESSION	BC004799		
VERSION	BC004799.1	GI:13435914	
KEYWORDS	MGC.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 (bases 1 to 2848)
TITLE	Strausberg, R.
JOURNAL	Direct Submission
	Submitted (21-MAR-2001) National Institutes of Health, Mammalian
	Gene Collection (MGC), Cancer Genomics Office, National Cancer
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
	USA
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk
	Email: cgaps-r@mail.nih.gov
	Tissue Procurement: Gilbert Smith, Ph.D.
	cDNA Library Preparation: Life Technologies, Inc.
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
	DNA Sequencing by: Baylor College of Medicine Human Genome
	Sequencing Center
	Center code: BCM-HGSC
	Web site: http://www.hgsc.bcm.tmc.edu/cdna/
	Contact: amg@bcm.tmc.edu
	Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
	Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
	Richards, S., Gibbs, R.A.
	Clone distribution: MGC clone distribution information can be found
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
	Series: IRAK Plate: 10 Row: 1 Column: 16
	This clone was selected for full length sequencing because it
	passed the following selection criteria: hexamer frequency ORF
	analysis.
FEATURES	Location/Qualifiers
source	1..2848
	/organism="Mus musculus"
	/db_xref="taxon:10090"
	/map="CZECH II"
	/clone="MGC:8136 IMAGE:3589474"
	/tissue_type="Mammary tumor metastatized to lung. Tumor
	arose spontaneously from a senescent normal mammary
	(clonal) outgrowth infected with the virus MMTV."
	/clone_lib="NCI CGAP_Lu29"
	/lab_host="DH10B"
	/note="vector: pcMV-SPORT6"
	167..1543
	/codon_start=1
	/product="Unknown (protein for MGC:8136)"
	/protein_id="AAH04799.1"
	/db_xref="GI:13435915"
	/translation="MYIDDLPIWGIVEADENGEDYLYWYKKLEIFGNGNRIVDVM
	TSBGKVLVPNTKIQMSYVKKSDYKDFEDKLDPSFFOHRHWFISFEMMV
	IFLVGLVSMILMRTLRKYARYSKEEMDDMDRLDGEYGWKQVHGDFRPSHPLIF
	SSLIGSGCQIFAVSLVIVIAMIEDLYTERGSMSTAIFYAATSPVNGVGGSLVAR
	QGRRWIKOMFIFAGVLIPAMVCGTAFINFIATYHASRAIPFGTWAVGCCIFEVIL
	PLNLVGTILGRNLGSGPNPCRYNAVPRPIPEKKWPEPAVIVCLGILDPGSIITEM
	YFIFTFWAKYIYYVGFPMMLVLVILICIVTCIVCTVCTYELLNADRYRWQWTSFLSRA
	STAIYVYMSFYIYFFKMKYGLFQTSFYFYMVAFSTALGIMCGAIFYMGTSFVVRK
	IYTNVKID"
BASE COUNT	819 a 489 c 595 g 945 t
ORIGIN	
Query Match	77.08; Score 1386.2; DB 10; Length 2848;
Best Local Similarity	91.7%; Pred. No. 1.5e-308;
Matches 1466; Conservative	0; Mismatches 133; Indels 0; Gaps 0;
QY	201 CTGGGAGAAGCACCCTCAAGGGGTTGAATTTAGTGGATTTAGTATTTAAATTTAAA 260
Db	1 CGGGGAGAAGCACCCTCAAGGGGTTGAATTTAGTGGATTTAGTATTTAAATTTAAA 70
QY	261 GATGATGTGATGCCAGCCACTTACTGTGAAATTTGATTAGATAAAGAAAAGAGAGATGCA 320
Db	71 GATGATGTGATGCCAGCCACTTACTGTGAAATTTGATTAGATAAAGAGAAGAGATGCG 130
QY	321 TTTCTATATGCCATAAANAATCATTTACTGGTACCAGATGTACATAGATGATTACCAATA 380
Db	131 TTTCTGTATGTATCAAGAATCACTATTGGTATCAGATGTACATAGATGACTTACCAATA 190

Qy 381 TGGGGTATTGTTGGTGAGGCTGATGAAATCGGAAAGATTACTATCTTTGGACCTATAAA 440  
Db 191 TGGGGTATTGTTGGTGAAGCAGATGAAATGGGAAAGATTACTATCTTTGGACCTACAAA 250  
Qy 441 AAACCTGAAATAGGTTTAAATGAAATCGAATTTGTTGATGTTTAAATCTAACCTAGTGAAGGA 500  
Db 251 AAACCTGAAATAGGTTTAAATGAAATCGAATTTGTTGATGTTTAAATCTAACCTAGTGAAGGA 310  
Qy 501 AAGGTGAACTGGTTTCAAATACTAAATCCAGATGTCATATTCAGTGAATGAAATGAAAGAAAG 560  
Db 311 AAGGTCAAGCTGGTTTCAAATACTAAATACAGATGTCATATTCAGTGAATGAAATGAAAGAAAG 370  
Qy 561 TCAGATGTAATTTGGAAGATGCAATTTGACAAATATCTTCGATCGCTCTTTTTCACACAT 620  
Db 371 TCGGATGTAATTTGGAAGATGCAATTCGATGAATATCTTCGATCCATCTTTTTCACACAT 430  
Qy 621 CGGATTCATTTGGTTTTCAAATTTTCAACTCCTTCATGATGGTGATCTTCTTGGTGGGCTTA 680  
Db 431 AGGATTCACCTGGTTTTCAAATTTTAAATTCCTTCATGATGGTGATCTTCTTAGTGGGATTA 490  
Qy 681 GTTTCATGATTTTAAATGAGAACATTAAGAAAGATTATGCTCGGTACAGTAAGAGAGAA 740  
Db 491 GTTTCATGATTTTAAATGAGAACATTTAAGGAAAGATTATGCCCGATACAGTAAAGAGAA 550  
Qy 741 GAAATGATGATATGGATAGAGACCTAGGAGATCAATATATGATGGAACACAGGTGCATGGA 800  
Db 551 GAAATGATGATGATGAGACACCTAGGAGACAGTATGGCTGGAAGCAGGTGCATGGA 610  
Qy 801 GATGATTTTGAACCATCAAGTCACCCACTGATATTTTCCCTCTCTGATGTTGTTCTCGGATG 860  
Db 611 GATGTTGTCAGCCGTCAGCTCAAGCTCACCTCTGATCTTCTCCCTCAATTTGGCTCTCGGATGT 670  
Qy 861 CAGATATTGCTGCTCTCATCGTTTATTATTTGTTCCAAATGATAGAGATTTATATACT 920  
Db 671 CAGATATTGCTGCTCTCATGTTTATTATTTGTTCCCATGATAGAGATTTATATACA 730  
Qy 921 GAGAGGGGATCAATGCTCAGTACAGCCATATTTGCTCTATGCTGCTAGCTTCTCCAGTGAAT 980  
Db 731 GAGAGGGGGTCAATGCTCAGCAGACCATATTTGCTCTATGCTGCTACATGCCAGTGAAC 790  
Qy 981 GGTTATTTTGGAGAGCTGTTATGCTAGACAGAGGAAAGAGATGGATTAAGACAGATG 1040  
Db 791 GGGTACTTTTGGAGGAGCTGTTATGCTAGACAGAGGAAAGAGATGGATTAAGACAGATG 850  
Qy 1041 TTTTATTTGGGGCATTCCTTATCCAGCTATGGTGTGGCAGCTGCCTTCTTCATCAATTTTC 1100  
Db 851 TTTTATCGGGGCATTCCTTATCCAGCTATGGTGTGGCAGCTGCATTTCTTCATCAACTTTT 910  
Qy 1101 ATAGCCATTTTATACCATGCTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCCGTTTGT 1160  
Db 911 ATAGCCATTTTATACCATGCTCTAGAGCCATTCCTTTTGGAAACAATGGTGGCCGTTTGT 970  
Qy 1161 TGCATCTGTTTTTGTATTTCTCTCTAAATCTTGTGGTACAATCTATGTCGCGGAAAT 1220  
Db 971 TGCATCTGTTTTTGTATTTCTCTCTAAATCTGCTGGTACAATCTATGTCGCGGAAAT 1030  
Qy 1221 CTGTCAGGTCAGCCCAACTTTCTGTCGTGTCGAATGCTGCTGCTATACCGGAG 1280  
Db 1031 CTGTCGGGTCAGCCCAACTTTCTGTCGTGTCGAATGCTGCTGCTATACCGGAG 1090  
Qy 1281 AAAAAATGGTTTATGAGGCTCGGGTTATTTGTTTGGCTGGGGTGGAAATTTTACCTTTTGGT 1340  
Db 1091 AAAAAATGGTTTATGAGGCTCGGATTTATTTGTTTGGCTGGGAGGAATTTTAACTTTTGA 1150  
Qy 1341 TCAATCTTTATGAAATGATTTTCATCTTTCACGCTCTTTCTGGGCAATATAGATCTATAT 1400  
Db 1151 TCAATCTTTATGAAATGATTTTCATCTTTCACGCTCTTCTGGGCAATATAGATCTACTAT 1210  
Qy 1401 GTCTATGGCTTCATGATGCTGGTGGCTGTTATCTCTGTCATTCGACTGCTGTCGTCAC 1460  
Db 1211 GTCTATGGCTTCATGATGCTGGTGGCTGTTATCTCTGTCATCTCTGTCATGCTGCTGTGACC 1270

Qy 1461 ATTGTGTGCACATATTTTCTACTAAATGCAGAAGATTACCGTGGCAATAGCAAGTTTT 1520  
Db 1271 ATTGTGTGCACATACTTCCTGCTAAACGCAGAGGATTACAGGTGGCAGTGGACGAGTTTC 1330  
Qy 1521 CTCTCTGCTGCATCAACTGCAATCTATGTTTACATGATTTCCCTTTTACTACTATTTTTC 1580  
Db 1331 CTGTCCTGGCGTCGCGCAATCTAGCTTTACATGTAATCTCTTTTACTACTACTCTCTTC 1390  
Qy 1581 AAACAAAGATGATGCTTATTTTCAACATCATTTTACTTTTGGATATATGGCGTATTT 1640  
Db 1391 AAACAAAGATGATGCTTATTTTCAACATCATTTTACTTTTGGCTACATGGCGGTGTT 1450  
Qy 1641 AGCACAGCCTTTGGGATAATGTTGGAGCATTGGTTTACATGGGAACAAGTGCCTTTGTC 1700  
Db 1451 AGCACAGCCTTTGGGATAATGTTGGAGCAATCGTTTACATGGGAACAAGTGCCTTTGTT 1510  
Qy 1701 CGAAAAATCTATCTAATGTGAAAAATTTGACTAGAGACCCAGAAAAACCTTGGAACTTTGGA 1760  
Db 1511 CGAAAAATCTATCTAATGTGAAAAATTTGACTAGATATCTGAGACAACATGGAACCTCGGA 1570  
Qy 1761 TCAATTTCTTTTTCATAGGGTGAACCTTGCACAGCAA 1799  
Db 1571 TCAGTTTCTCTTTTCATAGGGTGAACCTTGCACAGAAAA 1609  
RESULT 10  
A78983  
LOCUS A78983 1317 bp DNA linear PAT 20-OCT-1999  
DEFINITION Sequence 2 from Patent EP0848059.  
ACCESSION A78983  
VERSION A78983.1 GI:6092106  
KEYWORDS  
SOURCE unidentified.  
ORGANISM unclassified.  
REFERENCE 1 (bases 1 to 1317)  
AUTHORS Lenzen,G. and Strosberg,A.D.  
TITLE MAMMALIAN ICYP (IODOCYANOPINDOLOL) RECEPTOR AND ITS APPLICATIONS  
JOURNAL Patent: EP 0848059-A 2 17-JUN-1998;  
VETIGEN (FR)  
FEATURES  
Location/Qualifiers  
source  
1..1317  
/organism="unidentified"  
/db\_xref="taxon:32644"  
BASE COUNT 355 a 211 c 288 g 463 t  
ORIGIN  
Query Match 73.2%; Score 1317; DB 6; Length 1317;  
Best Local Similarity 100.0%; Pred. No. 1.2e-292;  
Matches 1317; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 357 ATGTACATAGATGATTTACCAATATGGGGTATTGTTGGTGAGGCTGATGAAATGGAGAA 416  
Db 1 ATGTACATAGATGATTTACCAATATGGGGTATTGTTGGTGAGGCTGATGAAATGGAGAA 60  
Qy 417 GATTACTATCTTTGGGACCTATAAAAACTTGAATAGTGTTTAATGGAATCGAATTTGTT 476  
Db 61 GATTACTATCTTTGGACCTATAAAAACTTGAATAGTGTTTAATGGAATCGAATTTGTT 120  
Qy 477 GATTGTTAATCTAATAGTGAAGGAAAGGTGAAACTGTTTCCAAATACTAAATCCAGATG 536  
Db 121 GATTGTTAATCTAATAGTGAAGGAAAGGTGAAACTGTTTCCAAATACTAAATCCAGATG 180  
Qy 537 TCATATTTCACTAAATGGAAGGAGTCAGATGTAATTTTCAAGATCGATTTGCAAAATAT 596  
Db 181 TCATATTTCACTAAATGGAAGGAGTCAGATGTAATTTTCAAGATCGATTTGCAAAATAT 240  
Qy 597 CTTGATCCGCTCCTTTTTCACATCGGATTCATTGGTGTTCATTTTCAATTTTCAACTCCTTCATG 656  
Db 241 CTTGATCCGCTCCTTTTTCACATCGGATTCATTGGTGTTCATTTTCAATTTTCAACTCCTTCATG 300  
Qy 657 ATCGTGATCTTCTTGGTGGGCTTAGTTCATGATTTTAAATGAGAACATTAAGAAAAAGAT 716  
Db 1 ATCGTGATCTTCTTGGTGGGCTTAGTTCATGATTTTAAATGAGAACATTAAGAAAAAGAT 716



|||||  
Db 781 TTTGGAACAATGTTGGCGTGTGTCATCTGTTTTTGTGTTATCTCTCTCTAAATCTT 840  
Qy 1197 GTTGGTACAATCTTGCCGGAATCTGTAGGTGAGCCCAACTTTTCTTGTGCGTCAAT 1256  
Db 841 GTTGGTACAATCTTGCCGGAATCTGTAGGTGAGCCCAACTTTTCTTGTGCGTCAAT 900  
Qy 1257 GCTGTGCTGCTCTATACCGGAGAAAAATGGTTTCATGAGCGCTGCGGTATTGTTTGC 1316  
Db 901 GCTGTGCTGCTCTATACCGGAGAAAAATGGTTTCATGAGCGCTGCGGTATTGTTTGC 960  
Qy 1317 CTGGGTGGAATTTTACCTTTTGGTTCATCTTTATTGAAATGTTATTCATCTTCACGTCT 1376  
Db 961 CTGGGTGGAATTTTACCTTTTGGTTCATCTTTATTGAAATGTTATTCATCTTCACGTCT 1020  
Qy 1377 TTCTGGGCATATAGATCTATTATGCTATGCGTTCATGATGCTGGTGGTTCATCTG 1436  
Db 1021 TTCTGGGCATATAGATCTATTATGCTATGCGTTCATGATGCTGGTGGTTCATCTG 1080  
Qy 1437 TGCATTGTGACTGCTCTGTGACTATTGTGTGCACATATTTTCTACTAAATGCAGAAGAT 1496  
Db 1081 TGCATTGTGACTGCTCTGTGACTATTGTGTGCACATATTTTCTACTAAATGCAGAAGAT 1140  
Qy 1497 TACCGGTGCAATGGACAAGTTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATG 1556  
Db 1141 TACCGGTGCAATGGACAAGTTTCTCTCTGCTGCATCAACTGCAATCTATGTTTACATG 1200  
Qy 1557 TATTCCTTTTACTACTATTTTTCAAAACAAGATGATGCTGCTTATTTCAACATCATTT 1616  
Db 1201 TATTCCTTTTACTACTATTTTTCAAAACAAGATGATGCTGCTTATTTCAACATCATTT 1260  
Qy 1617 TACTTTGCAATATGCGCGTATTATTAGCACAGCCTTGGGGATAATGCTGGAGCGATT 1673  
Db 1261 TACTTTGCAATATGCGCGTATTATTAGCACAGCCTTGGGGATAATGCTGGAGCGATT 1317  
  
RESULT 12  
AK027438  
LOCUS  
DEFINITION Homo sapiens cDNA FLJ14532 fis, clone NT2RM2000402, weakly similar to ENDOSOMAL P24A PROTEIN PRECURSOR.  
ACCESSION AK027438  
VERSION AK027438.1 GI:14042111  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens teratocarcinoma cell line: NT2 cDNA to mRNA, clone lib: NT2RM2 clone: NT2RM2000402.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Shiratori, A., Sudo, H., Wagatsuma, M., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Chiba, Y., Ishida, S., Murakawa, K., Ono, Y., Takiguchi, S., Watanabe, S., Kimura, K., Murakami, K., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y., Ninomiya, K. and Iwayanagi, T.  
NEDO human cDNA sequencing project  
Unpublished  
2 (bases 1 to 2311)  
Isogai, T. and Otsuki, T.  
Direct Submission  
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.  
Location/Qualifiers

source 1. .2311  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="NT2RM2000402"  
/cell\_line="NT2"  
/cell\_type="teratocarcinoma"  
/clone\_lib="NT2RM2"  
/note="cloning vector: pUC19FL3-mRNA from uninduced NT2 neuronal precursor cells."  
48. .1037  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAB55110.1"  
/db\_xref="GI:14042112"  
/translation="MDDMRDLGDEYKQVHGVDVFRPSHPLIFSSLTSGSGCQIFAV SLIVIVAMIEDLYTERGMSLSTAIFYAATSPVNGYFGSLARGGRRWIKOMFIG AFLIPAMVCGTAFINFIAYIHASRAIPFGTMVAVGCCICFFVILPNLVGTLGRNL SQQNPFCRVNAVPRPIPEKFMPEPAVIVCLGILPFGSIFTEMYFIFTSPWAYFIY YVGFEMMLVLVICIVTCVTIVCTYFLINAEDYRWQWTSFLSAASTAIYVVSFY YPFKTKMYGLFQTSFYFGYMAVFSTALGIMCGAIGMGTSAFVRKIYTNVKID"  
BASE COUNT 655 a 382 c 458 g 816 t  
ORIGIN  
Query Match 61.3%; Score 1104; DB 9; Length 2311;  
Best Local Similarity 100.0%; Pred. No. 1.4e-243;  
Matches 1104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 697 TGAGAACATTAAGAAAAGATTATGCTCGGTACAGTAAGAGGAGAAATGGATGATATGG 756  
Db 1 TGAGAACATTTAAGAAAAGATTATGCTCGGTACAGTAAGAGGAGAAATGGATGATATGG 60  
Qy 757 ATAGAGACCTAGAGAGTAATATGGATGGAACAGGTGCATGGAGATGATATTTAGACCAT 816  
Db 61 ATAGAGACCTAGAGAGTAATATGGATGGAACAGGTGCATGGAGATGATATTTAGACCAT 120  
Qy 817 CAAGTCACCCACTGATATTTTCTCTCTGATTTGGTTCTGGATGTCAGATATTTGCTGTGT 876  
Db 121 CAAGTCACCCACTGATATTTTCTCTCTGATTTGGTTCTGGATGTCAGATATTTGCTGTGT 180  
Qy 877 CTCTCATCGTTATTATTTGGCAATAGTAGAAGATTATATCTAGAGAGGGATCAATGC 936  
Db 181 CTCTCATCGTTATTATTTGGCAATAGTAGAAGATTATATCTAGAGAGGGATCAATGC 240  
Qy 937 TCAGTACAGCCATATTTGCTCTATGCTGCTACGCTCCAGTGAATGGTATTTTGGAGGAA 996  
Db 241 TCAGTACAGCCATATTTGCTCTATGCTGCTACGCTCCAGTGAATGGTATTTTGGAGGAA 300  
Qy 997 GTCTGTATGCTAGACAAGGAGGAGATGGATAAAGCAGATGTTTATTTGGGSCATTC 1056  
Db 301 GTCTGTATGCTAGACAAGGAGGAGATGGATAAAGCAGATGTTTATTTGGGSCATTC 360  
Qy 1057 TTATCCAGCTATGGTGTGGGCACCTGCCCTTCATCAATTTTCATAGCCATTTATTTACC 1116  
Db 361 TTATCCAGCTATGGTGTGGGCACCTGCCCTTCATCAATTTTCATAGCCATTTATTTACC 420  
Qy 1117 ATGCTTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCGGTTTGTGTCATCTGTTTTTG 1176  
Db 421 ATGCTTTCAAGAGCCATTCCTTTTGGAAACAATGGTGGCGGTTTGTGTCATCTGTTTTTG 480  
Qy 1177 TTATTTCTCTCTAAATCTTTGTTGGTACAAATCTTGCGCGAAATCTGTTCAGTCAGCCCA 1236  
Db 481 TTATTTCTCTCTAAATCTTTGTTGGTACAAATCTTGCGCGAAATCTGTTCAGTCAGCCCA 540  
Qy 1237 ACCTTCTCTCTGTCGTCGTCGCTGCTGCTTACCTATACCGGAGAAAAATGCTTCATGG 1296  
Db 541 ACCTTCTCTCTGTCGTCGTCGTCGCTGCTTACCTATACCGGAGAAAAATGCTTCATGG 600  
Qy 1297 AGCTGCGGTTATGTTGCTCGGTGGAATTTTACCTTTTGCTTCAATCTTTATTTGAAA 1356  
Db 601 AGCTGCGGTTATGTTGCTCGGTGGAATTTTACCTTTTGCTTCAATCTTTATTTGAAA 660  
Qy 1357 TGATTTCATCTTCACGTCTTTCTGGGCATATTAAGATCTATTTATGCTATPGCTTCATGA 1416  
|||||

Db	661	TGATTTTCATCTTCACGGCTTTCCTGGGCAATATAAGATCTATTATGTCTATGGCTTCATGA	720
Qy	1417	TGCTGGTGCTGGTTATCTCTGTGCATTTGTGACTGTCTGTGTGACTATTTGTGTGCACATATT	1476
Db	721	TGCTGGTGCTGGTTATCTCTGTGCATTTGTGACTGTCTGTGTGACTATTTGTGTGCACATATT	780
Qy	1477	TTCTACTAAATGCAGAGATTTACCGTGGCAATGGACAAGTTTCTCTCTCTGCATGCAAA	1536
Db	781	TTCTACTAAATGCAGAGATTTACCGTGGCAATGGACAAGTTTCTCTCTCTGCATGCAAA	840
Qy	1537	CTGCAACTCTATGTTTACATGTATTCCTTTTACTACTATTTTTCAAAACAAGATGTATG	1596
Db	841	CTGCAACTCTATGTTTACATGTATTCCTTTTACTACTATTTTTCAAAACAAGATGTATG	900
Qy	1597	GCTTATTTCAAACATCATTTTACTTTTGATATATGGCGGTATTTTAGCACACGCTTGGGGA	1656
Db	901	GCTTATTTCAAACATCATTTTACTTTTGATATATGGCGGTATTTTAGCACACGCTTGGGGA	960
Qy	1657	TAATGTGTGGAGCGATGGTTTACATGGGAACAAGTGCCCTTTGTCGCAAAAATCTATACTA	1716
Db	961	TAATGTGTGGAGCGATGGTTTACATGGGAACAAGTGCCCTTTGTCGCAAAAATCTATACTA	1020
Qy	1717	ATGTGAAATTTGACTAGACAGCCCAAGAAACCTGGAACTTTTGGATCAATTTCTTTTTCAT	1776
Db	1021	ATGTGAAATTTGACTAGACAGCCCAAGAAACCTGGAACTTTTGGATCAATTTCTTTTTCAT	1080
Qy	1777	AGGGGTGCAACTTGCACAGCAAAA	1800
Db	1081	AGGGGTGCAACTTGCACAGCAAAA	1104
RESULT 13			
A78984			
LOCUS	A78984	965 bp	DNA linear PAT 20-OCT-1999
DEFINITION	Sequence 3 from Patent EP0848059.		
ACCESSION	A78984		
VERSION	A78984.1 GI:6092107		
KEYWORDS	unidentified.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 965)		
AUTHORS	Lenzen G. and Strosberg, A. D.		
TITLE	MAHMALIAN ICYP (10DOCTANOPINDOLOL) RECEPTOR AND ITS APPLICATIONS		
JOURNAL	Patent: EP 0848059-A 3 17-JUN-1998;		
FEATURES	VETIGEN (FR)		
source	Location/Qualifiers		
	1. .965		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
BASE COUNT	246 a	163 c	214 g 342 t
ORIGIN			

Query Match	53.5%	Score 963.4	DB 6	Length 965
Best Local Similarity	99.9%	Pred. No. 3.1e-211		
Matches 964	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy 531	CAGATGTCATATTTCAGTAAATGCAAAAGTCAGATGTGAATTTGAAGATCGATTTCAC	590		
Db 1	CAGATGTCATATTTCAGTAAATGCAAAAGTCAGATGTGAATTTGAAGATCGATTTCAC	60		
Qy 591	AAATATCTTGATCCGTCCTTTTTTCAACATCGGATTCATTCGGTTTTCAAATTTTCAACTCC	650		
Db 61	AAATATCTTGATCCGTCCTTTTTTCAACATCGGATTCATTCGGTTTTCAAATTTTCAACTCC	120		
Qy 651	TTCATGATGGTGATCTCTCTCGTGGGCTTACGTTTCATGATTTTAAATGAGAACATTAA	710		
Db 121	TTCATGATGGTGATCTCTCTCGTGGGCTTACGTTTCATGATTTTAAATGAGAACATTAA	180		
Qy 711	AAAGATTTATGCTCGGTACAGTAAAGAGAGAAATGAGATGGATAGAGACCTAGGA	770		
Db 181	AAAGATTTATGCTCGGTACAGTAAAGAGAGAAATGAGATGGATAGAGACCTAGGA	240		

QY	771	GATGAATATGGATGGAACAGAGTGCATGGAGATGTATTTAGACCATCAAGTCACCCACTG	830
DB	241	GATGAATATGGATGGAACAGAGTGCATGGAGATGTATTTAGACCATCAAGTCACCCACTG	300
QY	831	ATATTTTCCCTCTCGATTGGTTCTCGATGTCAGATATTTGCTGCTCTCTCATCGTTATT	890
DB	301	ATATTTTCCCTCTCTCGATTGGTTCTCGATGTCAGATATTTGCTGCTCTCTCATCGTTATT	360
QY	891	ATTGTGTCGAATGATAGAAGATTTATATACTGAGAGGGGATCAATGCTCAGTACAGCCATA	950
DB	361	ATTGTGTCGAATGATAGAAGATTTATATACTGAGAGGGGATCAATGCTCAGTACAGCCATA	420
QY	951	TTTGCTATGCTGCTACGCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTAGA	1010
DB	421	TTTGCTATGCTGCTACGCTCCAGTGAATGGTTATTTTGGAGGAAGTCTGTATGCTAGA	480
QY	1011	CAAGGAGGAAGGAGATGATAAAGCAGATGTTATTGGGGCATTCCTTTATCCCAGCTATG	1070
DB	481	CAAGGAGGAAGGAGATGATAAAGCAGATGTTATTGGGGCATTCCTTTATCCCAGCTATG	540
QY	1071	GTGTGTGCACCTGCCCTTCTTCATCAATTTCATAGCCATTTATTACCATGCTTCAAGAGCC	1130
DB	541	GTGTGTGCACCTGCCCTTCTTCATCAATTTCATAGCCATTTATTACCATGCTTCAAGAGCC	600
QY	1131	ATTCCTTTTGGAAACAATGGTGGCCGTTTGTGTCATCTGTTTTTGTGTTATTCTTCCCTCTA	1190
DB	601	ATTCCTTTTGGAAACAATGGTGGCCGTTTGTGTCATCTGTTTTTGTGTTATTCTTCCCTCTA	660
QY	1191	AATCTTGTGGTACAAATACTTTGGCGGAATCTGTCAAGTCAGCCAGCCAACTTTCCTTGTGCT	1250
DB	661	AATCTTGTGGTACAAATACTTTGGCGGAATCTGTCAAGTCAGCCAGCCAACTTTCCTTGTGCT	720
QY	1251	GTCATGCTGTGCCCTGCTCTATACCGGAGAAAAATGGTTTCATGGAGCCGCGGTTATT	1310
DB	721	GTCATGCTGTGCCCTGCTCTATACCGGAGAAAAATGGTTTCATGGAGCCGCGGTTATT	780
QY	1311	GTTTGCCCTGGGTGGAATTTTACCTTTTGGTTCAATCTTTATTGAAATGTATTTTCATCTTC	1370
DB	781	GTTTGCCCTGGGTGGAATTTTACCTTTTGGTTCAATCTTTATTGAAATGTATTTTCATCTTC	840
QY	1371	ACGCTTTTCTGGGCATATAAAGATCTATTATGTCATGCTTCATGATGCTGGTGGTGGTT	1430
DB	841	ACGCTTTTCTGGGCATATAAAGATCTATTATGTCATGCTTCATGATGCTGGTGGTGGTT	900
QY	1431	ATCCTGTGCATTTGTGACGTGCTGTGTGACTATTGTGTGCACATATTTTCTACTAAATGCA	1490
DB	901	ATCCTGTGCATTTGTGACGTGCTGTGTGACTATTGTGTGCACATATTTTCTACTAAATGCA	960
QY	1491	GAAGA 1495	
DB	961	GAAGA 965	
RESULT	14		
LOCUS	A91327		
DEFINITION	Sequence 3 from Patent WO9826065.	965 bp	DNA
ACCESSION	A91327		
VERSION	A91327.1	GI:6740330	
KEYWORDS	unidentified.		
SOURCE	unidentified.		
ORGANISM	unidentified.		
REFERENCE	1 (bases 1 to 965)		
AUTHORS	Lenzen,G. and Strosberg,A.D.		
TITLE	MAMMALIAN ICY (IODOCYANOPINDOLOL)		
JOURNAL	Patent: WO 9826065-A 3 18-JUN-1998;		
FEATURES	VEITIGEN (FR); LENZEN GERLINDE (FR)		
	Location/Qualifiers		
source	1..965		
	/organism="unidentified"		
	/db_xref="taxon:32644"		
BASE COUNT	246 a 163 c 214 q 342 t		







BASE COUNT	672 a	651 c	613 g	726 t
ORIGIN				

Query Match 42.9%; Score 771.4; DB 3; Length 2662;  
Best Local Similarity 65.7%; Pred. No. 5.1e-167;  
Matches 1123; Conservative 0; Mismatches 586; Indels 0;

Qy	31	TGCCCCGACCCGGCGGACGACGACACACAGTATCAAGATAAAGAGGAAGTTGTCT	90
Db	254		91
		TGTGCTCTCCGAGGAGATGACGACCAATCAAAAGTACAAATGACCGGAGAGGTGGTAC	313
Qy	91	TATGATGAATACTGTGTGGCCCTACCAATAATCGTCAAGAAACATATAGTACTTTTCCAC	150
			151
Db	314	TGTGATGAACCGTGGGCCGTACCAANATCGCGAGGAGAGTACGCGTACTTCTCTC	373
			374
Qy	151	TTCCATTCTGTGTGGGTCAAAAAAAGTATCAGTCAITTTACCAAGAACTCTGGGAGAG	210
			211
Db	374	TCCCTTTTTCAGTGGCCAGAACTCCGTCAATATCGCACTACCAACGAGACGCTGACGGAGG	433
			434
Qy	211	CAC TTC AAGGGT TGAAT TGGAA TTTAG TGGT TGGAT ATTAAT TTAAGATGATGTGA	270
			271
Db	434	CGTCTCAAGGAGTCGAGCTAGAGTTCACTGGCTACGAGATGGAGTTCAAGACGACGCC	493
			494
Qy	271	TGCCAGCCACTTACTGTGAATTTGATTTAGATAAAGAAAAGAGAGATGCATTTGTATATG	330
			331
Db	494	CCAAATCGGTCTATCTGATGGTCACCTTGCAGGAGGAGACGCCAAGGCATTCACCTATG	553
			554
Qy	331	CCATAAAATCATTAATGTTACCAATGTACATAGATCAITTTACCAATATGGGATATTG	390
			391
Db	554	CCGTGAAGAACAGTACTGGTACCAATGTACATCATGATGGACTTGCCCATTTGGGAAAAG	613
			614
Qy	391	TTGGTAGGCTGATGAAATGGAGAAGATTACTATCTTTGGACCTATAAAAAATTTGAAA	450
			451
Db	614	TCGGTAGCGCGACGACGCGGATGGCAAGTACTATATCTTCACGCACAGAAGTTCGACA	673
			674
Qy	451	TAGTTTTTATGAAATCGAATTTGTTGATGTTTAATCTAACTAGTGAAGAAAGGTGAAC	510
			511
Db	674	TCGGCTACAAATGCCAGCAAACTGTTGGATATCACCTGACCAAGGAGGCCGCGAGAAC	733
			734
Qy	511	TGGTTCCAATACTAAATCCAGATGTATATTCAGTAAATTTGAAAAAGTCAGATGTGA	570
			571
Db	734	TCAAGCGGGATCCCACATCACTTCTCTACGAGTCACTTGAAGCCGACAGGTGG	793
			794
Qy	571	AATTTGAAGATCGATTGTGACAAATATCTTGATCCGTCCTTTTTTCAACATCGGATTCATT	630
			631
Db	794	AGTTCAGAATCGATTTCGACAAAGTACGTGGATCCCAACTTCTTCACGACAGGATCCACT	853
			631
Qy	631	GGTTTTCAITTTTCAACTCCTTCATGATGGTGCATCTCTTGTTGGGCTAGTTTTCATGA	690
			631
Db	854	GGTTCAGCATCTTCAACAGTCTCATGATGGTGCATCTCTCTGGTGGGCTGTTGTCCATGA	913
			691
Qy	691	TTTTAATGAGAACATTAAGAAAAGATTATGCTCGGTACAGTAAAGAGGAAGAAATGGATG	750
			914
Db	914	TTCTGATGCGAACTCTGCGCAAGGATATTGCTCGGTACAGTAAAGCAGGAGAAATCGACG	973
			751
Qy	751	ATATGATAGACACTAGGAGATGAATGGAATGGAACAGTGCATGGATGTATTTA	810
			974
Db	974	ACATGAGCGAGATCTTGTGTATGAATACGGCTGGAAGCAGGTGATGGCATGTCTTCC	1033
			811
Qy	811	GACCATCAAGTCAACCACATATTTTCCCTCTCTGATTGTTCTGGATGTCAAGATATTG	870
			1034
Db	1034	GTCTCCGCCCAACACACTGCTCTCTCGCGCTTGGTGGCGCTGATACCAACTGATTT	1093
			871
Qy	871	CTGTGCTCTCATCGTTATTATTGTTGCAATGATAGAAGATTTATATCTGAGAGGGAT	930
			1094
Db	1094	CGGTGTATTTCTGTGTGATCATGTTGCCATAGTTGGTGAATTTGACAGGAACCGGCT	1153
			931
Qy	931	CAATGCTCAGTACAGCCATTTGTCTATGCTGCTACGCTCCACGTGAATGTTGTTT	990
			1154
Db	1154	CCATGCTGCCCGGTATATTGTGTATGCGGCCACCTCACCAATCAATGGATACTTTG	1213
			991
Qy	991	GAGGAAGTCTGTATGTAGACAAGGAGGAGAGATGGATAAAGCAGATGTTTATTTGGG	1050

Search completed: January 20, 2003, 14:10:57  
Job time : 4585 secs

Search completed: January 20, 2003, 14:10:57  
Job time : 4585 secs



GenCore version 5.1.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2003, 14:15:05 ; Search time 50 seconds  
(without alignments)  
1107.469 Million cell updates/sec

Title: US-09-319-724A-14  
Perfect score: 3089  
Sequence: 1 AALLMLLLLPRTADEHEH.....ICYMGTSFAVRKIVTNVKID 576

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1668	54.0	589	2 D86243	hypothetical prote
2	1146.5	37.1	592	2 D84431	probable endosomal
3	918	29.7	656	2 T28058	hypothetical prote
4	898	29.1	659	2 D84633	probable multisp
5	854	27.6	637	2 H96592	probable multisp
6	847.5	27.4	655	2 T32472	hypothetical prote
7	843	27.3	639	2 T50793	hypothetical prote
8	694	22.5	667	2 T64915	EMP70 protein prec
9	688	22.3	629	2 T39285	probable transmem
10	667.5	21.6	672	2 S2673	probable membrane
11	650	21.0	527	2 H85135	hypothetical prote
12	447	14.5	260	2 D86280	protein F10B6.2 [i
13	438	14.2	706	2 S0616	hypothetical prote
14	121	3.9	605	2 H95240	conserved hypothet
15	121	3.9	605	2 A99705	hypothetical prote
16	119	3.9	578	1 F64578	oligoendopeptidase
17	117.5	3.8	417	2 F97789	ampg protein limpo
18	117.5	3.8	820	2 A40204	Na+/H+-exchanging
19	115	3.7	502	1 I30010	NADH2 dehydrogenas
20	115	3.7	528	2 S05063	hypothetical prote
21	115	3.7	578	2 B71934	oligoendopeptidase - H
22	114	3.7	478	2 T11318	NADH2 dehydrogenas
23	113	3.6	570	2 S07744	NADH2 dehydrogenas
24	112.5	3.6	792	2 T06624	hypothetical prote
25	112	3.6	792	2 G84830	probable potassium
26	111	3.6	666	2 F58931	NADH2 dehydrogenas
27	110.5	3.6	2848	2 T32550	hypothetical prote
28	110	3.6	297	2 S24333	H+-transporting tw
29	110	3.6	818	2 A48858	Na+/H+-exchanging

RESULT 1

D86243

hypothetical protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: D86243

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ausen, N.F.; Hughes, B.; Hulzar, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86243

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-589 <STO>

A:Cross-references: GB:AE005172; NID:g1931647; PIDN:AAB65482.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08

Query Match 54.0%; Score 1668; DB 2; Length 589;

Best Local Similarity 54.2%; Pred. No. 9.8e-122;

Matches 313; Conservative 107; Mismatches 151; Indels 6; Gaps 4;

QY 5 LLLLLLPRTADEHEHTYQDKKEVVLWMTVGPYHNRQETKYKFSLPFCVSGSKKSISHYH 64

Db 14 LILVSLTPTFASDSOHKYQAEQVTLWNKVGYPNNPQETNYSLPFCRPGSNVY-HKW 72

QY 65 ETLGEALQCVLEFSGDLTKFKDDVPATYCEITDLDEKRDFAVYAKNHYQWYIDDL 124

Db 73 GGLGEVLGNNELIDSEIAIKFNKNVRSVICPLELDEAKVHKFKAIESYFEEFMGMF 132

QY 125 PTWIVGE--ADENGE--YLLWTYKKLEIGFNGRIVDVNLTSECKVKLVPTNTKIOMSY 180

Db 133 HVCCFVGEHLPHDKNSGKHVLYTHKNIVKYNKDQIHNLTQDNPRPLEAGKGMKDLTY 192

QY 181 SVYKKKSDVKFEDRDYLDFFOHRTHWFSTIENSFMVIELGLVSLMILMRLTKDYA 240

Db 193 SVQVPTPTNFARDEVYLDYFFEHQIHWFSIENSFMVIELGLVSLMILMRLNDYIA 252

QY 241 RYSKE-EEMDDMDRLDGLDEYGNKQVHGDFRPSHPHFSSLISSGCOIFAVSLIIVIA 299

Db 253 KYAREDDDDLESLESDVSESGNKLHVGDFRFPASSLVLLSAVVGTAQALALLVLLVMA 312

QY 300 MIEDLYTERGSMLSAIFVYAATSPVNGYFGGSLYARQGRRWIKOMFIGAFLIPAMVCG 359

hypothetical prote  
hypothetical prote  
hypothetical prote  
sulphakinin recepto  
ubiquinol-cytochro  
unknown protein T1  
ubiquinol-cytochro  
amino acid transpo  
MDR-type permease  
hypothetical prote  
hypothetical prote  
ubiquinol-cytochro  
sodium channel mna  
oligopeptide ABC t  
probable efflux pr

Db 313 IVGTLVGRGAIVTTTIVCYALTSTSVSGYVSGGMYSRSGGKHWTICMVLTLASLFFFLFCG 372  
Qy 360 TAFFNTAIYVHASRAJPFPGTMVAVCCICPFVILPLNLVGTILGRNLSSGPNPCRYNA 419  
Db 373 IGFLNTAIYVGSAAJPFPGTMVAVVFWIFGIFSLALLGTVVGRNNSGAPNPCRVT 432  
Qy 420 VPRPIPKKWFMEPAVIVCLGILPFGSIFITFIEMFYFTSWAYKIYVYVGMMLVLVILC 479  
Db 433 IPRPIPKKWLTPSVSLMGLLPFGSIFITFIEMFYFTSWAYKIYVYVGMMLVFLVILV 492  
Qy 480 IVTVCTIVCTVYFLNAEDYRWQWTSFSAASTAIYVYVYFYFFTKTKYGLPQTSFY 539  
Db 493 IVTVCTIVCTVYFLNAEDYRWQWTSFSAASTAIYVYVYFYFFTKTKYGLPQTSFY 552  
Qy 540 FGYMAVFSTALGIMCGAIGYMGCTSAFVRKIYTNVKID 576  
Db 553 FGYTMFCLGILCGAVYLGSLNLFVRRIYRNKCD 589  
RESULT 2  
D84431  
probable endosomal protein [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: D84431  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayan, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: AB4420; MUID:20083487; PMID:10617197  
A:Accession: D84431  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-592 <STO>  
A:Map position: 2  
A:Gene: At2g01970  
C:Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08  
Query Match 37.1%; Score 1146.5; DB 2; Length 592;  
Best Local Similarity 36.9%; Pred. No. 3.3e-81;  
Matches 214; Conservative 125; Mismatches 214; Indels 27; Gaps 10;  
Qy 14 RADEHTYQDKKEVVLNMTVGPYHNQETKYFSLPFCVSGKKSISHYHETLGEALQ 73  
Db 23 RSDASDRYKQGSVPLANKVGPYHNQETKYFSLPFCVSGKKSISHYHETLGEALQ 80  
Qy 74 VELEFSGLDIKFKDDVMPATYCEIDLKREKDAFYAIKNHYQMYIDDLPIWGIVGEA 133  
Db 81 DRLVSAPYKLNPRDEKDESEVYCKKLSREEVEHFRRAVEKDYFYQMYDDLPPIWGIVGEA 140  
Qy 134 DENGED-----YYLWTKKLEIGNGRIVDVN--LISEGKVKLVPTNKT--QMSYSVK 183  
Db 141 DKESKSDPSEFYKLYKIQHIFLYKQKRVINARMPSHSLVOLTEKDEDAEFMYTVK 200  
Qy 184 WKSDVKFEDRDKYLDPSFFOH--RIHWFSLFNFMVIFLVGLVSMILMRTLKRDYAR 241  
Db 201 WKETSTSEKRMKDYKAMSSSLPHHLEIHWFSILNSCVVLLTGLFATILMRVLKNDPMK 260  
Qy 242 YSKEEMDDMDRLDGEDYKQVGHGDFRPSHPLIFSLSLGSGCQIFAVSLIIVAMI 301  
Db 261 YAQDEAAD-DQ---EETGWKYIHGDVFRFPKNSLFAASLGSQGLTTLTIFIMLSIV 316  
Qy 302 EDLYT-ERGSMLSTAIYVAAATSPVNGYFGGSLYARQGRRWIKOMFTGAFILPAMVCGT 360  
Db 317 GVFPYNRGALFTALVVIYALTSGIATSSAFYQCLEGNWRNL---LTTGLFLFCGP 372  
Qy 361 AF----FINFIAIYHASRAJPFPGTMVAVCCICPFVILPLNLVGTILGRNLSSGPNPCRYNA 416  
Db 373 LFLTCFLNTVAIYASATAALDPFGIIVIVLWTLVTSFLVLVGLGIAGKNSKAEPQAPVR 432

Qy 417 VNAYPRPIPEKKWFMEPAVIVCLGILPFGSIFITFIEMFYFTSWAYKIYVYVGMMLVLV 476  
Db 433 TTKYPREIPLPWVRSRSPVQAMAGFLPFSAIYIELYIFASVWGHRIYTIYSILFIVFI 492  
Qy 477 ILCIVTVCTIVCTVYFLNAEDYRWQWTSFSAASTAIYVYVYFYFFTKTKYGLPQTSFY 536  
Db 493 ILLIVTAFITVATLYFQAAEDHEWNRSLFCGGSTGLFIYAYCLYIYIYARSMSGFMQT 552  
Qy 537 SFYFCYMAVFSTALGIMCGAIGYMGCTSAFVRKIYTNVKID 576  
Db 553 SFFPGYMACICYGFLMLGTGVFRAALLFVRHIYRSIKCE 592  
RESULT 3  
T28058  
hypothetical protein ZK858.6 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C:Accession: T28058  
R:White, S.  
submitted to the EMBL Data Library, September 1996  
A:Reference number: Z20462  
A:Accession: T28058  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-656 <WIL>  
A:Cross-references: EMBL:Z79759; PIDN:CA802141.1; GSPDB:GN00019; CESP:ZK858.6  
A:Experimental source: clone ZK858  
A:Genetics:  
A:Gene: CESP:ZK858.6  
A:Map position: 1  
A:Introns: 1/3; 36/3; 78/3; 165/2; 222/2; 309/1; 377/1; 537/3; 596/1; 623/2  
C:Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08  
Query Match 29.7%; Score 918; DB 2; Length 656;  
Best Local Similarity 33.9%; Pred. No. 2.1e-63;  
Matches 195; Conservative 110; Mismatches 218; Indels 52; Gaps 9;  
Qy 45 KYFSLPFCVSGKKSISHYHETLGEALQVELEFSGLDIKFKDDVMPATYCEIDLKREK 104  
Db 91 FEYYSVPFCKPANGDIQYKSENLEVMRGDRIVTTPYAFHMKNEQCVCVCSNLSKENV 150  
Qy 105 DAFVIAIKHNYQMYIDDLPIWGIVGEADENGEDYIYLTWTKKLEIGNGN----- 155  
Db 151 ALFERIRIQEYSAHLIVDNLPVATVINPA-QSGDVYVYDLYGRLGWIGDNKAVFLNNHQLF 209  
Qy 156 -----RIVDVNLTSEGKVKLVN-TK-----IQMSYSVK 183  
Db 210 VVKYHQHTPGLYRVVGFVRPRSRISATKNDGTCMPDGGKHVELGDSQSVDFSYVA 269  
Qy 184 WKSDVKFEDRDKYLDPSFFQHRHWFSLFNFMVIFLVGLVSMILMRTLKRDYARYS 243  
Db 270 FEESDVPWASRWDVYLTTKAVD--IHWFSILNSIVVLSLGSFVSVTIVTRVRRDIAQYN 327  
Qy 244 KEEMDDMDRLDGEDYKQVGHGDFRPSHPLIFSSSLGSGCQIFAVSLIIVAMIED 303  
Db 328 RDDEED----DTLEETGWLKLVHGDVFRPPHQMILVNNMVGTIQLLGNLSATVVCAMLCGM 383  
Qy 304 LY-TERGSMSTAIYVAAATSPVNGYFGGSLYARQGRRWIKOMFTGAFILPAMVCGTAF 362  
Db 384 LSPASRGSLMAAANVPLFCFMGLISGYHAGRIYKTMKGNRPIRCAVQTATLPSLILGAF 443  
Qy 363 FINFIAIYHASRAJPFPGTMVAVCCICPFVILPLNLVGTILGRNLSSGPN-PPCRVNAV 421  
Db 444 LLNFLICKHSSGAVPFCMTVALLIMWFCIDMPLIFLGFYFG--YRKQYPHPVTRTNQIP 501  
Qy 422 RPIPEKKWFMEPAVIVCLGILPFGSIFITFIEMFYFTSWAYKIYVYVGMMLVLVILCIV 481  
Db 502 RQVPEQPFLLIPSLSLIAGVLPFGAMFIELFFINATWENQFYFLFGFLFIVSLILAS 561  
Qy 482 TVCVCTIVCTVYFLNAEDYRWQWTSFSAASTAIYVYVYFYFFTKTKYGLPQTSFY 541  
Db 562 TQISVVAITYSLCAENYRWNRSRFVSGSFYVMAVAFYNTKLTIEGFPVTLVFS 621



T32472

hypothetical protein F08F1.7 - *Caenorhabditis elegans*C:Species: *Caenorhabditis elegans*

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 20-Jun-2000

C:Accession: T32472

R:Fulton, B. submitted to the EMBL Data Library, September 1997

A:Description: The sequence of *C. elegans* cosmid F08F1.

A:Reference number: Z21174

A:Accession: T32472

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-655 &lt;F08F1&gt;

A:Cross-references: EMBL:AF026213; PIDN:AAB71307.1; GSPDB:GN00028; CESP:F08F1.7

A:Experimental source: strain Bristol N2; clone F08F1

C:Genetics:

A:Gene: CESP:F08F1.7

A:Map position: X

A:Introns: 42/3; 65/2; 96/3; 183/3; 376/1; 539/2; 612/3

C:Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08

Query Match 27.4%; Score 847.5; DB 2; Length 655;

Best Local Similarity 31.3%; Pred. No. 6.3e-58; Matches 189; Conservative 111; Mismatches 224; Indels 79; Gaps 11;

QY 45 YKFSLPFCVCGSKSIHSHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDL----- 99

Db 61 YEYHSEDFCTVNEDESP--VENLGOVLFGRIIRPSYKVAFLQBEKRLVCDTKKARGS 118

QY 100 --DKKRDAFYAIKNHWYQMYIDLLPI-----WGIVEADENG----- 137

Db 119 ADDLAKRLQLQAMTLYQHHWIVDNNPVTFCEFKNLQNMDCVTTGFPVGCVDVEQGYQHD 178

QY 138 -----EDYLYWTKYKKLEIGN-----GNRIVDNL----- 162

Db 179 ACVLNOKYKTPNNFYIFNVDVIEIYRDNTDGNFLEHKVGGRIIRIDVAPRSIKSSSS 238

QY 163 ---TSEGVKVLVPTNK-----IQMSYSVKWKSDVDFEDFKYLDPSFPQHRHWFISF 214

Db 239 SLDCSDSAEPIADAKSSEAEITYSYKWKTKDIKWSRWD-YILQSMPTNIQWFSIM 297

QY 215 NSPMWVFLVGLYSMLMRTLKDYARYSKEEEMDDMDRLGDEYQKQVHGDFVRPSH 274

Db 298 NSLVIVFLTGMVGMIIIMRTLHRDIDRYNRL----DTEEDAQEEFGWKLVHGDFVTRPY 353

QY 275 PLIFSSLGSGCOIFAVSLVIVIAMIEDLY--TERGSMSTAIFFVAATSPVNGYFGGSL 333

Db 354 PMLLSVFIGAGCOTLLMVSVTLVPAFLGFLSPANRGLITFAFFVYLVGIVAGYISARL 413

QY 334 YARQGGRRWIKOMETGAFILPAMVCGTAFINFIAYIYHASRAIPFGTMVAVCCICFFVI 393

Db 414 YKTFEGHWNKTNLVMTAFVPGILFTIFFPSNTLLWTKGSSAAVPGTLLVLLVLFIS 473

QY 394 LPNLVGTILGRNLGQNPFCRVNAVPRPIPEKKWMEPAVIVCLGILPFGSIFIEWY 453

Db 474 VPMTFVAGYFGFKRG--IEAPVTRNKIPROVPEQTEYTKPLPGLMGGILPFGICIQLF 532

QY 454 FIFTSFWAYKIYYVYGMVLVILCVTVCTIVCTYVCTYVCTYVCTYVCTYVCTYVCTY 513

Db 533 FILNSIWAHOTYYFMGFLVLYLILITSEATILLAYHLCAEDYHWMRSFMTSGFTA 592

QY 514 IYVYMSFYFFPKTKMYGLFQTSFYFGYMAVFSTALGIMCAIGYMGTSFAVRKIYTNV 573

Db 593 IYLFYCIHFFNKKLISGVIISILYFSYISFVFMFLMTGCTIGFLATYFYVRKIYGSV 652

QY 574 KID 576

Db 653 KVD 655

RESULT 7

T50793

hypothetical protein T30N20\_110 - *Arabidopsis thaliana*C:Species: *Arabidopsis thaliana* (mouse-ear cress)

C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 28-Jul-2000

C:Accession: T50793

R:Revan, M.; Peters, S.A.; van Staveren, M.; Dirkse, W.; Stiekema, W.; Bancroft, I.;

submitted to the Protein Sequence Database, July 2000

A:Reference number: Z25240

A:Accession: T50793

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-639 &lt;BEV&gt;

A:Cross-references: EMBL:AL365234

A:Experimental source: cultivar Columbia; BAC clone T30N20

C:Genetics:

A:Map position: 5

A:Introns: 46/3; 101/3; 133/2; 170/3; 208/2; 254/3

A:Note: T30N20\_110

C:Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08

Query Match 27.3%; Score 843; DB 2; Length 639;

Best Local Similarity 34.0%; Pred. No. 1.4e-57; Matches 201; Conservative 104; Mismatches 216; Indels 70; Gaps 13;

QY 42 QETKYFSLPFCVCGSKSIHSHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLK 101

Db 63 QLPYSYSLPFCRPSK--IVDSTENLGEVLGRDRIENAPYSAQ-----MCNILGRVTLD 115

QY 102 EKRDFAFYAIKNHWYQMYIDLLPIWGIIVEADENG--YYLWTKYKKLEIGN-- 155

Db 116 KTAAPKEKIDDEYRWNMLDNLPLVVPVIERVDQGSFVYQVGVHGLKQVGEKREQ 175

QY 156 -----DVLNLTSEG-----KVKLV---PNTKIOMS 179

Db 176 FFMHNLAFTRYHRDIQTDAAIRIVGFVKPYSKHIEGESEKRLTCTDPTKRLVV 235

QY 180 YS-----VKKK-----SDVKFEDRDKYLDPSFQHRHWFISFMMVIFLVGL 226

Db 236 SSATPQVEVEQKKEIIFTYDESEVKWASRWDYLLMS--DNQIHWFISVNSLMIVLFSGM 293

QY 227 VSMILMRTLKDYARYSKEEEMDDMDRLGDEYQKQVHGDFVRPSHPLIFSSLGSGC 286

Db 294 VAMIMLRTLKDYARYSKEEEMDDMDRLGDEYQKQVHGDFVRPSHPLIFSSLGSGC 349

QY 287 QIFAVSLVIVIAMIEDLY--TERGSMSTAIFFVAATSPVNGYFGGSLYARQGGRRWIK 345

Db 350 QCLGMHVFMTIFAMGLFSLSPNRGLMTAMLLWVPMGLFAGYASSRLYKMKFGTEWKRI 409

QY 346 MFTGAFILPAMVCGTAFINFIAYIYHASRAIPFGTMVAVCCICFVILPLNLVGTILGR 405

Db 410 APTAFELFAVVSATFVNLALIWGQKSGAVPFGTMTALIFLWFGISVPLVFGYIGF 469

QY 406 NLSGQPNFCRVNAVPRPIPEKKWMEPAVIVCLGILPFGSIFIEWYFIFTSFWAYKIY 465

Db 470 KRPAADD--PVKTNKIDPRQIPQOAWYNNPVFSLIGILPFGAVFIEELFILLSIMLNQFY 528

QY 466 YVYGMVLVILCVTVCTIVCTYVCTYVCTYVCTYVCTYVCTYVCTYVCTYVCTYVCTY 525

Db 529 YIFGFLFVFLIVTCAETVTVLCYFQLCSEYDLWMMRSYLTSGSSALYLPLATYFF 588

QY 526 FTKTKYGLFQTSFYFGYMAVFSTALGIMCAIGYMGTSAPVRKIYTNVKID 576

Db 589 TKLIQITKLVSAMLYFGYMLIASYAFVFLTGTIGFYACILWFTRLIYSSVKID 639

RESULT 8

S64915

EMP70 protein precursor - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein L2385; protein YLR083c

C:Species: *Saccharomyces cerevisiae*

C:Date: 01-Aug-1995 #sequence\_revision 24-May-1996 #text\_change 20-Jun-2000

C:Accession: S64915; S25110; A47165

R:Pohl, T.M. submitted to the Protein Sequence Database, May 1996

A:Reference number: S64899



Db 515 HPYLFMEFGSFFCGILGVLTTTCIMVSIITVYFQLCSENNWNWRSFIPGFCGIYVIFSV 574  
QY 522 YVYFFTKMYGLFQTSFYFGYMAVESFALGIMCGAIGYMGTSFAFVKIYTNVKID 576  
Db 575 FYWFFKISSSLATAVLYFGYLLISLVFLGSGVGFAGFLVKNKIYASIKID 629  
RESULT 10  
S52673  
probable membrane protein YDR107c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein YD9727.03c  
C:Species: Saccharomyces cerevisiae  
C:Date: 19-May-1995 #sequence\_revision 01-Sep-1995 #text\_change 19-Apr-2002  
C:Accession: S52673  
R:Murphy, L.; Shore, L.; Harris, D.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: S52671  
A:Accession: S52673  
A:Molecule type: DNA  
A:Residues: 1-672 <MUR>  
A:Cross-references: EMBL:Z48758; NID:g747879; PID:g747882; GSPDB:GN000004; MIPS:YDR107c  
C:Genetics:  
A:Gene: MIPS:YDR107c  
A:Cross-references: SGD:S0002514  
A:Map position: 4R  
C:Superfamily: Schizosaccharomyces pombe probable transmembrane protein SPBC1105.08  
C:Keywords: transmembrane protein  
F:311-327/Domain: transmembrane #status predicted <TM1>  
F:384-400/Domain: transmembrane #status predicted <TM2>  
F:408-424/Domain: transmembrane #status predicted <TM3>  
F:444-460/Domain: transmembrane #status predicted <TM4>  
F:479-495/Domain: transmembrane #status predicted <TM5>  
F:536-552/Domain: transmembrane #status predicted <TM6>  
F:561-577/Domain: transmembrane #status predicted <TM7>  
F:618-634/Domain: transmembrane #status predicted <TM8>  
F:640-656/Domain: transmembrane #status predicted <TM9>  
Query Match 21.6%; Score 667.5; DB 2; Length 672;  
Best Local Similarity 25.8%; Pred. No. 6.3e-44;  
Matches 169; Conservative 125; Mismatches 252; Indels 109; Gaps 16;  
QY 21 TYQDKKEVLMNTVGP--YHNHQE-----TYKFSLPFCVGSKKKSISHY 63  
Db 28 TYHSGDEIPLLVKNLTPSIYHQDEBNDVSGDKHEFLYSYDNKRFHFCRPEHVEKQ 87  
QY 64 HETLGEALQGVLEFSLDKIKFKDDVNPATYCEIDLDKERDAFVAIKNHYWQMYIDD 123  
Db 88 PESLGSVIFGDIYNSPQLNMLEEKCEKVALCKSTIPGKDAKFINTLIKSGFFQNMLVDG 147  
QY 124 LP-----TWG-----IVGEA-----DENGEDY----- 140  
Db 148 LPAARKAYDSRTKTYNYGTGFGELGFTDVKQTVDGKAVPTMEELTSEASNEVDILDAFLP 207  
QY 141 -----YLWYKKLEIGN--GN-----RIVDVNL-----TSEGV 168  
Db 208 KNVKNLVLKTVLPEYFVNHFDIEVEFHDRGNDRYVGVIVNPNVSTERSPPGACSTTGK 267  
QY 169 KLPNTK---IQMSYSVKWKKSDVKEDRPDKYLDPSFFOHRTHWFSIFNSFMWVFLVG 225  
Db 268 LIIDEDKDNVEYFYSVKFVASDTVWATRDYKYL--HIYDPQIQWFSLNFVSVILLSS 325  
QY 226 LVSMILMRLTKRDYARYSKKEEMDDMDRLGDYGNKQVHGDVFRPSSHPLIFSSLIGSG 285  
Db 326 VMHSLRLALKSLARYNEL----NLDNEFHEDSGWKLGHGDVFRTPSKSMLLSILVSGS 381  
QY 286 QITFAV---SLIIVIAMIEDLYTERGSMSTAIFVAATSPVNGVFGGSLARQGRRW 342  
Db 382 MQLFLVMVCSIFFAAVGLVSPV--SRGSLPTVMFVLVYALFGVGSYASMGVYKFFRGPYW 439  
QY 343 IKOMFTGAFILIPAWCGTAFFINFIAIYVHASRAIPGTMVAVCCICFFVLLPLNLVGTI 402  
Db 440 KANNIITPLLGAIFLLIYIMNFFLLFAHSSGVIPARSLFFILLWFLVSPVLSFAGSI 499

QY 403 LGRNLGQPNFPCRVNAVPRPIPEKKWFMEPAVIVCLGGILPFGSIFIEFYFTSEWAY 462  
Db 500 VAHKQCNDWDEHPTKTNQIARQIPYQPWYLRTAQATLIAGIFSGSIARELYFYVSSLWFN 559  
QY 463 KIYVYGVFMMVLVLILICIVTCVTVTYTFLNAEDYRWQWTSF--LSAASATAIYVYMSF 521  
Db 560 KIFMFGFLFLSLLLTATLSLTILITYSLGLENLWQWRSFIITGGLCSTYTFIHSI 619  
QY 522 YVYFFTKMYGLFQTSFYFGYMAVESFALGIMCGAIGYMGTSFAFVKIYTNVKID 576  
Db 620 --LFTKFKLGVTIVVLYLGYSLIISALCCVVTGAIGFFSSMFFIRKIYSAIRKE 672  
RESULT 11  
H85135  
hypothetical protein AT4g12650 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: H85135  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Reference number: A85001; MUID:20083488; PMID:10617198  
A:Accession: H85135  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-527 <STO>  
A:Cross-references: GB:NC\_001268; NID:g7267967; PIDN:CAB78308.1; GSPDB:GN00140  
C:Genetics:  
A:Gene: AT4g12650  
A:Map position: 4  
Query Match 21.0%; Score 650; DB 2; Length 527;  
Best Local Similarity 32.8%; Pred. No. 1.1e-42;  
Matches 160; Conservative 91; Mismatches 171; Indels 66; Gaps 17;  
QY 128 GIYGEADENGEDYLLWYKKLEIGN--GNRIY--DVNLTFSEGVKVL-----VPNT---- 174  
Db 67 GVISEAD-----KKKALGYEIVGVFEVPCSVKYDAEKMTKLHMYDVPVSNCP 115  
QY 175 -----KIOMSYSVKWKKSDVKEDRPDKYLDPSFFOHRTHWFSIFNSFMWVIF 222  
Db 116 ELDKAQIIEKHERITFTYEVEFYKSETRWP SRWDVYL--KMEGARVHWFSLNSLAWIFF 173  
QY 223 LVGLVSMILMRLTKRDYARYSKKEEMD----DMDRLGDYGNKQVHGDVFRPSSHPLI 277  
Db 174 LAGIVFVIFLRTVRRDLTKY---EELDKAQAOQMEELS---GWKLVGVDVFRPEMSKL 227  
QY 278 FSSLIGSGCQIFAVSLIIVIAMIEDLY--TERGSMSTAIFVYAATSPVNGYFGGSLYAR 336  
Db 228 LCIMVGDGVRIITGMVAVTIVFAALGFMSPASRGMLLTGMILYFLGIVAGYAGVRLWRT 287  
QY 337 QGG--RRWIKQMFIGAFLIPAMVCGTAFFI-----NFIATYYHASRAIPFGTMTAVCCICF 390  
Db 288 VKGTSEGRWSLSIACFFP---GIAFVLLTVLNFLLSSNSTGAIPISLYFELLALWF 343  
QY 391 FVTLPLNLVGTIILGRNLGQPNFPCRVNAVPRPIPEKKWFMEPAVIVCLG--GILPGSIF 449  
Db 344 CISVPLTLFGFGUGTAAEA--IQPVRTNQIPREIPERKY---PSWLLVLGAGILPFGTLF 399  
QY 450 IEMYFIFTSEWAYKIYVYGVFMMVLVLILICIVTCVTVTYTFLNAEDYRWQWTSFSLA 509  
Db 400 IELFFIFSSILWGREYVYVFGFGLIVLLVLLVVCVAEVSVLTYMHLVCVEDWRWNWKAIFYAS 459  
QY 510 ASTAIYVYMSFYFYFFKTK--MYGLFQTSFYFGYMAVESFALGIMCGAIGYMGTSFAVRK 568  
Db 460 GSVALLYFAIYSINLYVFDLQSLGSPVSAIMLYIGYSLMAITAMLATGTIGFLTFSFYFVHY 519  
QY 569 IYTNVKID 576  
Db 520 LFSNVKID 527





```
QY 244 KEEMDDMDRLDGEYGMKQVHGDVFRPSHPLI-----FSSLI-----GSG 285
Db 55 KNEIDLI-----GFERFRYRIVPPVLMVLVMTPTFLVRQDYVAGIG 99
QY 286 CQI-----FAVSLIIVIIV-----MIEDLYTERGSM-----STAFVYAATSPV 325
Db 100 GQTAGVLGFMNTFYELLTGGSYESQIPHLFVHNWSLAVEHYIILWGLAVWFLSKQAKS 159
QY 326 NGYFGGSLYAROGGRWIK--OMFIGAFLIPAMVCGTAFFINFIAIYYHASRAIPP--GT 381
Db 160 NGQLKGMVFLLSAVAFILSFFSMFIGSLVTSY--SSVYFSSLTHVY-----PFFLGS 210
QY 382 MVAVCCICFFVILPLNLVG---TILGRNLSQPNPCRVNAVPRPIPEKWMFMEPAVIV 437
Db 211 MLA-----TIVGVRQTSLVKQL-----DKIWLRLKTLVV 240
QY 438 CLGGILPFGSIFIEYFI-FTSFNAYKIYYVYGFMMVLVILCIYVCTIVCTYFLLNA 496
Db 241 FGGG---FGFLVLLTFFVKFTYLFAYLI---GFLASLAALAMILAA-----RVLHE 286
QY 497 EDIRWQWT---SFLSAATAIYVYMSFYFFBKTK-----MYGLFQTSFY 539
Db 287 KTHHIQESKIISFLADTSYAVILFHWHPFYIIFSOLTSNLLAVLLTLCISYGFASLSFY 344

RESULT 15
A99705
hypothetical protein spr1868 [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: A99705
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; E
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:2142945; PMID:11544234
A:Accession: A99705
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-605 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00670.1; PID:gl5459558; GSPDB:GN00174
C:Genetics:
A:Gene: spr1868

Query Match 3.9%; Score 121; DB 2; Length 605;
Best Local Similarity 19.9%; Pred. NO. 0.16;
Matches 83; Conservative 59; Mismatches 116; Indels 160; Gaps 22;

QY 207 RTHWFSIFN-----SPMMVILFVG-----LVSMLMRTLKDYARYS 243
Db 2 RIKWFSIRIIGLLVLYHFFQTFPGGFGVDVFFTFSGFLITALLIE-----EFS 54
QY 244 KEEMDDMDRLDGEYGMKQVHGDVFRPSHPLI-----FSSLI-----GSG 285
Db 55 KNEIDLI-----GFERFRYRIVPPVLMVLVMTPTFLVRQDYVAGIG 99
QY 286 CQI-----FAVSLIIVIIV-----MIEDLYTERGSM-----STAFVYAATSPV 325
Db 100 GQTAGVLGFMNTFYELLTGGSYESQIPHLFVHNWSLAVEHYIILWGLAVWFLSKQAKS 159
QY 326 NGYFGGSLYAROGGRWIK--OMFIGAFLIPAMVCGTAFFINFIAIYYHASRAIPP--GT 381
Db 160 NGQLKGMVFLLSAVAFILSFFSMFIGSLVTSY--SSVYFSSLTHVY-----PFFLGS 210
QY 382 MVAVCCICFFVILPLNLVG---TILGRNLSQPNPCRVNAVPRPIPEKWMFMEPAVIV 437
Db 211 MLA-----TIVGVRQTSLVKQL-----DKIWLRLKTLVV 240
QY 438 CLGGILPFGSIFIEYFI-FTSFNAYKIYYVYGFMMVLVILCIYVCTIVCTYFLLNA 496
Db 241 FGGG---FGFLVLLTFFVKFTYLFAYLI---GFLASLAALAMILAA-----RVLHE 286
```

```
QY 497 EDIRWQWT---SFLSAATAIYVYMSFYFFBKTK-----MYGLFQTSFY 539
Db 287 KTHHIQESKIISFLADTSYAVILFHWHPFYIIFSOLTSNLLAVLLTLCISYGFASLSFY 344

Search completed: January 20, 2003, 14:36:44
Job time : 54 secs
```

Result No.	Query			Length	DB	ID	Description
	Score	Match					
1	950.5	30.8	642	9	US-10-201-964-1		Sequence 1, Appli
2	674	21.8	125	10	US-09-915-582-52		Sequence 52, Appli
3	674	21.8	125	10	US-09-915-582-68		Sequence 68, Appli
4	130.5	4.2	351	10	US-09-820-893-114		Sequence 114, Appli
5	119.5	3.9	58	10	US-09-864-761-34915		Sequence 34915, A
6	119.5	3.9	58	10	US-09-864-761-36367		Sequence 36367, A
7	117.5	3.8	307	10	US-09-820-893-117		Sequence 119, Appli
8	117.5	3.8	322	10	US-09-820-893-117		Sequence 117, Appli
9	114	3.7	234	9	US-10-001-857-176		Sequence 176, App
10	108.5	3.5	433	9	US-09-738-626-6716		Sequence 6716, Ap
11	106	3.4	1681	10	US-09-920-653-3		Sequence 3, Appli
12	104.5	3.4	445	9	US-09-738-626-3627		Sequence 3627, Ap
13	104.5	3.4	822	10	US-09-824-734-3		Sequence 3, Appli
14	104	3.4	385	9	US-09-738-626-4559		Sequence 4559, Ap
15	104	3.4	791	9	US-09-738-626-4530		Sequence 4530, Ap
16	102	3.3	609	10	US-09-801-368-50		Sequence 50, Appli
17	101.5	3.3	305	10	US-09-815-242-13291		Sequence 13291, A
18	101.5	3.3	305	10	US-09-815-242-13553		Sequence 13553, A
19	101.5	3.3	3092	10	US-09-801-368-172		Sequence 172, Appli



Db 121 LYLQ--ROY-----IPVKMSKAFWIESWEYAMMY-----VGSLVVILCLS 160  
QY 390 FVILPLNLVGTILGRNLG--OPN-----FPC--RVN---AVPR 422  
Db 161 FFLSSWDFIPAVYGFLSVDPDTPNIGLWYFFAEMFEHSLFFVCVFOINVFFYTIPL 220  
QY 423 PIPEKK-----WFMEPAVIVCLGILPFGSIFIEYFIPTSFWAYKIYVYVGFMLVLVI 477  
Db 221 AIKLKHPHIFEMFIQIAVIAIFKSYPTVGD--VALYMAFFPVW---NHLRYELRNIEVL 274  
QY 478 LCIVTVCVTIVCHYFLLNAEDYRWQTSFLSAASATYI-----VVMYGFYFFFK 527  
Db 275 TCIIIVCSLFFPLVHL-----WYIPGNANSFEYAITLTFNVGQILLISDIYFAFL 326  
QY 528 TKMY----GLFQTS 537  
Db 327 RREYLYTHGLYLA 340

RESULT 5  
US-09-864-761-34915  
; Sequence 34915, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 34915

; LENGTH: 58  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL049539.16  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.1  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.5  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2  
; OTHER INFORMATION: EST\_HUMAN HIT: AW372226.1, EVALUATE 1.00e-30  
; OTHER INFORMATION: SWISSPROT HIT: P32802, EVALUATE 2.00e-05  
US-09-864-761-34915

Query Match 3.9%; Score 119.5; DB 10; Length 58;  
Best Local Similarity 43.6%; Pred. No. 0.00042;  
Matches 24; Conservative 9; Mismatches 21; Indels 1; Gaps 1;  
QY 377 IPFGTWAVCCICOFFVILPLNLVGTILGRNLGSGQNPFCRVNAVPRPIPEKKWFM 431  
Db 1 VPPEPTVALLCMWFGISLPLVLGYVFGFRKQPYDN-PVRTNQIPIROIPEQRWYM 54  
RESULT 6  
US-09-864-761-36367  
; Sequence 36367, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US 60/180,312  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687

Query Match	3.98;	Score 117.5;	DB 10;	Length 307;
Best Local Similarity	23.5%;	Pred. No. 0.0054;		
Matches	58;	Conservative 38;	Mismatches 80;	Indels 71; Gaps 14;

  

QY	277	IFSSLIGSGCOIFAVSLIVIIAMIEDLYTERGMSLTALFV----	YAATSPVNGYEGGS	332
		:     :     :	:     :     :	
Db	67	ILSCVAKSTCAI--NNTLIARFI-----LTTIKSAGFLSAIFLALAT	YOSLYPLTLFVPGL	120
		:     :     :	:     :     :	
QY	333	LYARQGGRRIKMFIGAFLIPAMVCCTAFPI---NFIAIYVHASRAIPFGTMAVOCIC	389	
		:     :     :	:     :     :	
Db	121	LYLQLQ--RQY-----IPVKMKSKAIFSWEYAMMY-----	VGSLVVIICLS	160
		:     :     :	:     :     :	
QY	390	FFVLPLNLWGTILGRNLSG----OPN-----	FPC--RVN-----	422
		:     :     :	:     :     :	
Db	161	FFLLSSMDGFIPAVYGFTLSVPDPTLPNTIGLEWYFAEMFEHFSLFVCGVFQINFFVFIPL	220	
		:     :     :	:     :     :	

```

RESULT 9
US-10-001-857-176
; Sequence 176, Application US/10001857
; Publication No. US20020193500A1
; GENERAL INFORMATION:
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Hevve
; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and
; FILE REFERENCE: DEX-0273
; CURRENT APPLICATION NUMBER: US/10/000
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/252,054
; PRIOR FILING DATE: 2000-11-20

```

	: NUMBER OF SEQ ID NOS: 208	
:	: SOFTWARE: PatentIn version 3.1	
:	: SEQ ID NO 176	
:	: LENGTH: 234	
:	: TYPE: PRT	
:	: ORGANISM: Homo sapien	
:	: US-10-001-857-176	
	Query Match	3.7%; Score 114; DB 9; Length 234;
	Best Local Similarity	19.2%; Pred. No. 0.0078;
	Matches	39; Conservative 34; Mismatches 90; Indels 40; Gaps 5;
Qy	358 CGTAFFI-NFTAIYHSAIRPCTMTAVCCICFCFVILPNLVGTILGRNLSCGPPCR	416
Dd	60 CGVFELFCVCVFCOCVLFSFLFFSLLCFEFVLFMSLVMVLCGL-----110	
Qy	417 VNAVPRPIPEKKMFEPAVICLGILLPGFSIFTEMFI---FTSWAYKIYYVGEM 472	
Dd	111 -----LFFFGVSLSCLFGFLFSFLSVFVVVLVFAFWFPACFCVFPF 159	
Qy	473 LVLVILCIVTIVCTYFLLNAEDRWQWTSLSAASATIAIYMY-----SFYYYYF 526	
Dd	160 CLLVFEFFFL-----FEVRLFPSSESLFSFEAFVPVLCLEFLPLVVFVEFFFFF 210	
Qy	527 KTKMYGLFQTSFYGYMAVESTA 549	
Dd	211 SFSPFLPFVFFVFFFSSFGSS 233	
	RESULT 10	
	US-09-738-626-6716	
:	: Sequence 6716, Application US/09738626	
:	: Publication No. US20020197605A1	
:	: GENERAL INFORMATION:	
:	: APPLICANT: NAKAGAWA, SATOSHI	
:	: APPLICANT: MIKOGUCHI, HIROSHI	
:	: APPLICANT: ANDO, SEIKO	
:	: APPLICANT: HAYASHI, MIKIRO	
:	: APPLICANT: OCHIAI, KEIKO	
:	: APPLICANT: YOKOI, HARUHIKO	
:	: APPLICANT: TATEISHI, NAOKO	
:	: APPLICANT: SENOH, AKIHIO	
:	: APPLICANT: IREDA, MASATO	
:	: APPLICANT: OKAZAKI, AKIO	
:	: TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES	
:	: FILE REFERENCE: 249-125	
:	: CURRENT APPLICATION NUMBER: US/09/738, 626	
:	: PRIOR FILING DATE: 2000-12-18	
:	: PRIOR APPLICATION NUMBER: JP 99/377484	
:	: PRIOR FILING DATE: 1999-12-16	
:	: PRIOR APPLICATION NUMBER: JP 00/159162	
:	: PRIOR FILING DATE: 2000-04-07	
:	: PRIOR APPLICATION NUMBER: JP 00/280988	
:	: PRIOR FILING DATE: 2000-08-03	
:	: NUMBER OF SEQ ID NOS: 7059	
:	: SOFTWARE: PatentIn ver. 3.0	
:	: SEQ ID NO 6716	
:	: LENGTH: 433	
:	: TYPE: PRT	
:	: ORGANISM: Corynebacterium glutamicum	
:	: US-09-738-626-6716	
	Query Match	3.5%; Score 108.5; DB 9; Length 433;
	Best Local Similarity	19.8%; Pred. No. 0.054;
	Matches	77; Conservative 55; Mismatches 144; Indels 113; Gaps 19;
Qy	276 LFSSSLIGSGCGIPANSLIVIIVAMI--EDLYTERGSMLSTAIFYVAATSPPNGYF---GG 331	
Dd	19 VLILSGSGVIEWDFLVITGAALVNKNYPSPGNEFLSTILAYASFSLTFEPRPG 77	
Qy	332 SLVAROG---GRWRWKOMFGAEILPAMVCGTAFFINFAT-----369	
Dd	78 VFIEHGIDRIGRK--KTFLTTLMGGGTVAGLDDPDYNAIGTWAGTPILLMFRLIQIGI 135	

```

Db   1352 --RVIHILRPGKPKVPHDMLP-LILALPALLNISLILFLVMFYIAIFGMYNFAIVYKKE 1400

Qy   468 -----YGFMMLVL 475
      :| ||
Db   1409 AGINDSVNFETFGSSMLCL 1427

RESULT 12
US-09-738-626-3627
; Sequence 3627, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 3627
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3627

Query Match          3.4%; Score 104.5; DB 9; Length 445;
Best Local Similarity 19.6%; Pred. No. 0.13;
Matches 57; Conservative 55; Mismatches 116; Indels 63; Gaps

Qy   197 KYLDPSFOHRIHWFSIFNSFMMVIFLVGLYSMLMRTRLRKYDIARYSKSEEEEMDDMDRLG 256
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   171 QYVNLTFFE--TLWVS-----LVLVVG--SLTALGVKERRGRHPVLVANPDVVQTLLG 220
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   257 DEYGRQVHGCVFRPSSHPLFISSLIGSGCQIFAVSLIVIITAMI-BDLYTERGSMLSTA 315
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   221 Q--GFKLRLND-----RRAREVTYIRINSIPYAMAVFPFSETDDLKWQLSWFLILT 272
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   316 IFVYAATSPVNGYFGSGLYAROGGRWIKQMFIGAFILPAWVCATGFINFIAIFYHASR 375
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   273 TVTIYANLPFNPF-F-GSFGDRHG---WARTFWGSGIGGAUTLALVYIFPMFGVQAQMSN 338
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   376 AIPFGTMVAWCICFCFVILPLNLVCTILGNLSG-QPNFPCRVNAVPRPIPEKKWFMEPA 434
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   329 GVVFEGTTIA-----AGALFEVSLAGFVPLSAIAVSLDPK-----HPG 365
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Qy   435 VIVC-----LGGILPGSFIDEMVFIETSFWAYKIYYVYGFMMILVLVILCI 480
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db   366 AAMATYNLGVGCAVAVGPLLVAVF-----HPLIGPTGILLVIAL 405
      :||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 13
US-09-824-734-3
; Sequence 3, Application US/09824734
; Patent No. US20020083485A1
; GENERAL INFORMATION:
; APPLICANT: ZHU, JIAN-KANG
; APPLICANT: SHI, HUAZHONG
; APPLICANT: ISHITANI, MANABU
```



```

: PRIOR APPLICATION NUMBER: JP 00/159162
: PRIOR FILING DATE: 2000-04-07
: PRIOR APPLICATION NUMBER: JP 00/280988
: PRIOR FILING DATE: 2000-08-03
: NUMBER OF SEQ ID NOS: 7059
: SOFTWARE: PatentIn ver. 3.0
: SEQ ID NO 4559
: LENGTH: 385
: TYPE: PRT
: ORGANISM: Corynebacterium glutamicum
US-09-738-626-4559

```

Query Match	3.4%;	Score 104;	DB 9;	Length 385;	
Best Local Similarity	20.9%;	Pred. No. 0.12;			
Matches	62;	Conservative 50;	Mismatches 118;	Indels 66;	Gaps

  

QY	210	WFSIFNSFMVIFVLGVLVSMILNRTLRKDIARYSKSEEDMDDRDLDGEYKQVHGQVDF	269
		:    : : :    : : :    : : :    : : :	
Db	82	FFSILSIAPAVLLGVSVITFLA-----SDSTEILNLVRDEVQY-----VP	123
QY	270	RPSHPL--IFSSLLSGCOIFAVSLIVIAMIEDLYTERGSMSTALFVVAATSPVNG	327
		:    : : :    : : :    : : :    : : :	
Db	124	EDOSHVVNGVIDSIAGS---AAGGVGAVGVITALT-----SSAIVRAFSRCAN-	171
QY	328	YFGGSLYAROGGRWTKQMFIGAFLIPAMVCGTAFENIAIYVHASRAIPGTMVAVCC	387
		: :            : :    : :    : :    : :	
Db	172	----AVYRGSEGTLLKRAMLLFLNLALLG-----IILVSWVLNETVLNMGIFAPTA-	222
QY	388	ICFFVILPLNLVGTILGRNLSGP-----NFFCRNAV-----PRPIPEKKWFM	431
		:    : :    : :    : :    : :	
Db	223	-----EPLHLTNVLSFLTDREMPIWIWYRFPVVGVLIMFVATLYYWAAPNAPKFKWL	276
QY	432	EPAVIVCLGGILPFGSIFTEMFTSFWAYKIYYVYGFMMVLVLILCIIVTCVTI	487
		: :    : :    : :    : :    : :	
Db	277	SLGSLFAIVGILLAG---VGLNFYETLFFAASVSYGAVGSLLAFTALWVFNICLLI	329

RESULT 15

```

US-09-738-626-4530
; Sequence 4530, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4530
; LENGTH: 791
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-4530

```

Query Match 3.4%; Score 104; DB 9; Length 791;  
Best Local Similarity 22.6%; Pred. No. 0.3;  
Matches 65; Conservative 39; Mismatches 115; Indels

Qy	225	GLVSMILMRLTRKDYAR	SK-----EEMDDMRDL-----GDYGNKQVHGVD-----	268
		:	:	
Db	160	GIISMTEDADSAMD	ISAEDEKVTNLTDEYDDGLTVVNGNVFGAAATSLD	219
		:	:	
Qy	269	-----FRPSHPLFISS	LLIGSCQIFAVYSLIIVAMTEDLTYGRSMLSTA	315
		:	:	
Db	220	LVAAVLVITFGSFI	AGMPLI-SAIIGVIGITGQLATAFTDSVNDMTPTL	278
		:	:	
Qy	316	IFVVAATSPVNGYFG	SGLYARQG-----GRRWIKOMFTGAFLIPAM	356
		:	:	
Db	279	VGIDYALFIVS-RFN	ELISQTDANLEPKELAERLTWPAAARAHAMGNAV	337
		:	:	
Qy	357	VCGTAPFINFIATY	IYHASRAIPGTMVAVCC-----ICFVETLPLNLTG	413
		:	:	
Db	338	FAGTTVLIALVAL---	SIINIPETVNAIAAATVAIAVLVLSFPAALLG	392
		:	:	
Qy	414	PCRVNA--VPRPIPEK	-----KW-----FMPEPAVICLGGILPFGSIF	450
Db	393	AARVPGPKVPDPEDEK	PTMGLKWLVRKMPVAYLLVGVVLLCAIAI	439
		:	:	

Search completed: January 20, 2003, 14:15:02  
Job time : 251 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 20, 2003, 10:32:24 ; Search time 37 seconds  
(without alignments)  
458.043 Million cell updates/sec

Title: US-09-319-724A-14

Perfect score: 3089

Sequence: 1 AALWLLLLLPRTRADEHEH.....IGYMGTSFAVRKIYTNVKID 576

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCRUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	950.5	30.8	625	4	US-08-959-004-10
2	843.5	27.3	663	4	US-08-959-004-5
3	694	22.5	667	4	US-08-959-004-11
4	118.5	3.8	496	4	US-09-134-001C-3703
5	112.5	3.6	237	4	US-09-134-001C-3057
6	107.5	3.5	408	2	US-08-742-440A-6
7	107	3.5	333	4	US-09-576-160B-6
8	104.5	3.4	453	1	US-08-439-131A-5
9	104.5	3.4	453	1	US-08-440-674-4
10	101.5	3.3	2938	5	PCT-US94-00198-3
11	99.5	3.2	616	1	US-08-149-100-2
12	97.5	3.2	470	2	US-08-724-394A-10
13	96	3.1	357	5	PCT-US95-07180-3
14	95	3.1	367	2	US-08-514-451A-7
15	94.5	3.1	425	1	US-08-313-553-13
16	94.5	3.1	425	3	US-08-767-993-13
17	94.5	3.1	503	4	US-09-134-001C-3096
18	94	3.0	360	4	US-08-875-573-20
19	94	3.0	360	4	US-09-232-878-2
20	94	3.0	360	4	US-09-045-583-55
21	94	3.0	360	4	US-09-534-185-55
22	94	3.0	1873	1	US-08-336-257A-7
23	93	3.0	434	2	US-08-677-049-7
24	92.5	3.0	521	4	US-09-134-001C-4290
25	92.5	3.0	844	4	US-09-422-936-51
26	92.5	3.0	886	4	US-09-422-936-77
27	92.5	3.0	892	4	US-09-422-936-75

Sequence 71, Appl  
Sequence 49, Appl  
Sequence 4037, Ap  
Sequence 23, Appl  
Sequence 40, Appl  
Sequence 3, Appl  
Sequence 3593, Ap  
Sequence 1, Appl  
Sequence 2, Appl  
Sequence 10, Appl  
Sequence 4, Appl  
Sequence 3, Appl  
Sequence 4, Appl  
Sequence 4, Appl  
Sequence 80, Appl  
Sequence 47, Appl  
Sequence 45, Appl  
Sequence 7, Appl

28 92.5 3.0 899 4 US-09-422-936-71  
29 92.5 3.0 961 4 US-09-422-936-49  
30 92.5 3.0 1076 4 US-09-134-001C-4037  
31 91.5 3.0 409 4 US-09-326-203A-23  
32 91 2.9 464 4 US-09-724-864-40  
33 91 2.9 606 3 US-08-620-077B-3  
34 90.5 2.9 486 4 US-09-134-001C-3593  
35 90 2.9 352 4 US-09-576-160B-1  
36 90 2.9 352 4 US-09-576-160B-2  
37 90 2.9 360 4 US-08-833-752-10  
38 90 2.9 473 1 US-08-439-131A-4  
39 90 2.9 473 1 US-08-440-674-3  
40 90 2.9 1240 3 US-08-930-996A-4  
41 90 2.9 1873 1 US-08-435-675B-4  
42 89.5 2.9 411 4 US-08-887-534A-80  
43 89.5 2.9 844 4 US-09-422-936-47  
44 89.5 2.9 960 4 US-09-422-936-45  
45 89 2.9 464 4 US-09-604-978-7

#### ALIGNMENTS

RESULT 1  
US-08-959-004-10  
; Sequence 10, Application US/08959004  
; Patent No. 6197543  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Yue, Henry  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; APPLICANT: Kaser, Matthew  
; TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/959,004  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0414 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 625 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1665777

US-08-959-004-10

Query Match 30.8%; Score 950.5; DB 4; Length 625;  
Best Local Similarity 35.7%; Pred. No. 6.6e-85;  
Matches 214; Conservative 100; Mismatches 212; Indels 73; Gaps 11;  
QY 42 QETKYKSLPFCVGSKKSHYHETLGEALOGVELEFSGLDIKFKDDVMPATYCE-----WGI-VG 131  
Db 36 QLPVEYSLPFCOPSK--ITYKAENLGEVLGRDVRTPFQVLMNSEKKCEVLCSQSNKP 93  
QY 97 IDLDKRDADFVAIKNHYWQYIDPLPI-----WGI-VG 131  
Db 94 VTLVEOSRLVAERITEDYVTHLADNLPVATRLLELYNSRDSDDKKKEKDVQFEHGVRLG 153  
QY 132 EAD-----ENGEDYILWYK--KLEIGFNGNRIVDNLVLTSEGVKVLV 171  
Db 154 FTDVWKIYLHNHLSFILIYHREDMEEDQHTYVRVREIPIQSIIRLEDLKADEKSSCTLP 213  
QY 172 PNT-----KIQMSYSVKKKSDVKPDRFDKYLDPSFFQHRHWFHSIFNSFM 218  
Db 214 EGTNSSPOEIDPTKENQLYFTYSVHWEESDIKWAASRWDYLTMSDQ--IHWFSIINSVV 271  
QY 219 MVFLVCLVSKMLMRLTKDYARYSKKEEMDDMDROLGDEYKQVHGDFRPSHPLIF 278  
Db 272 VWFPLSGILSMIITIRLKDIANYNKEDDIE---DTMEESGKWLHGDVFRPPQYPMIL 327  
QY 279 SSLIGSGCQIFAVSLIIVIAMIEDLY-TERGSMSTAFIYVAATSPVNGYFGGSLYARQ 337  
Db 328 SSLGSGIQFCMLIVFVAMLGMLSPSRGALMTTACFLPMFGVGGFSAGRLYRTL 387  
QY 338 GGRWIKOMF IGAFILPAMYCGTAFFINFATYIYHASRAIPFGTMAVCCICFVILPLN 397  
Db 388 KGRWKKGAFCTATLPVGVVFGICFVLNCFIWKHSGGAVPPTMVALLCMWFGISLPLV 447  
QY 398 LVGTILGRNLSGQNPFCRVNAPRPIPEKKWFEPNAVIVCLGILPFGSIFTEMFIET 457  
Db 448 YLGYGFRKQPDYD-VPRTNQIPROIQEQRWYMNFRVGMILMAGILPFGAMFIELPFIS 506  
QY 458 SFWAYKIYVYVGMVLVLVILCTIVTCVTIVCTYFLLNAEDYRWQMTSFLSAASTAIYVY 517  
Db 507 AIWENQFYLLFGFLFVLVILVSCSQISIVMYVFOLCADIEDYRWWHRNFLVSGSAFYVL 566  
QY 518 MYSEYFFTKMYGLFQTSFYCYGMAVFTALGIMCGAIGYMGTSFAVRKIYTNVKID 576  
Db 567 VYAIYFVFNKLDIVERTPSLLYFGYALMVSFLLTGTIGTFYAAAYMVRKIYAAVKID 625

RESULT 2

US-08-959-004-5  
Sequence 5, Application US/08959004  
Patent No. 6197543

GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Yue, Henry  
APPLICANT: Corley, Neil C.  
APPLICANT: Lal, Preeti  
APPLICANT: Shah, Purvi  
APPLICANT: Kaser, Matthew  
TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/959,004  
FILING DATE: Herewith  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0414 US  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 663 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: ADRETUT06  
CLONE: 2822412  
US-08-959-004-5

Query Match 27.3%; Score 843.5; DB 4; Length 663;  
Best Local Similarity 31.2%; Pred. No. 2.4e-74;  
Matches 199; Conservative 120; Mismatches 234; Indels 85; Gaps 16;  
QY 9 LIPRTRADHEHTYQDKEEVVLMNMTVGPVHNQETKYFSLPFCVGSKKSHYHETLG 68  
Db 41 LAPVNCDEEKKDECKAEIELEFVNRILDSVES-VLPEYVTAQDFCOASEG--KRPSENLG 97  
QY 69 EALQGVLEEFSGLDIKFKDD-----VMPATY-CEILDREKRDADFVAIKNHYWQYID 122  
Db 98 QVLFGRIEIPSPYKFTFNKKETCKLYCTTKTYHTEKAEDKQKLEFKSKMLLNQHHWID 157  
QY 123 DLPI-W-----GIVGEADENGED-YVLT----- 144  
Db 158 NMPVTWCYDVEDQRCNPGFPIGCVITDKGHAKDACVSSDFHERDTYIFNHDVIKIY 217  
QY 145 YKKLETGFNGNRIV-----DYNLTSEGVKVLVNPNTKTMYSY 180  
Db 218 YHVETGSMGARLVAALKLEPKSFKHTHIDKPCSGPMDISNKSAGEI-----KIATYV 271  
QY 181 SVKWKSD-VKPEDRFKYLDPSPFOHRHWFISFNFMVIFLVLGSLMRLTKDY 239  
Db 272 SVSFEEDDKIRWASRDYILESMPTH-IQWFSIMNSLVIVLFLSGMVAIMLRLTKDI 330  
QY 240 ARYSKEEEMDDMDRLGDEYGNKQVHGVFRPSSHPILFSSLIGSGCQIFAVSLIIVIA 299  
Db 331 ARYN---QMDSTE-DAQEEFGKWLHGDIFRPRKGMILSVFLGSGTQILIMTFVTLFFA 386  
QY 300 MIEDLY-TERGSMSTAFIYVAATSPVNGYFGGSLYAROGGRWIKOMF IGAFILPAMYC 358  
Db 387 CLGFLSPANRGAALMTCAVVLWLLGTPAGYVAARFYKSFGEKKTNLLTSLCPGIVF 446  
QY 359 GTAFFINFATYIYHASRAIPFGTMAVCCICFVILPLNLVGTILGRNLSGQNPFCRVN 418  
Db 447 ADFFIMNLILWGEESAAIPFGTLVAIALALWFCISVPLTFIAGYFGKKNAIEH-EVRTN 505  
QY 419 AVPRPIPEKKWFMEPAVIVCLGILPFGSIFTEMFIETSEWAYKIYVYVGMVLVIL 478  
Db 506 QIPROIQEFSFYTKPLPGIIMGILPFGCIGFQLFLLINSHOWMYWFGFLVLFIL 565  
QY 479 CIVTVCTIVCTYFLLNAEDYRWQMTSFLSAASTAIYVYMYSPYFFTKMYGLFQTSF 538  
Db 566 VITCSEATILLCYFHLCAEDYHQRWRSFLTSGTAVYFLIYAVHYFYSKLQITGTASTIL 625  
QY 539 YFGYMAVFTALGIMCGAIGYMGTSFAVRKIYTNVKID 576  
Db 626 YFGYTMVIMVILFLFTGTIGTFACFWEVTKIYSVVKVD 663

```

RESULT 3
US-08-959-004-11
: Sequence 11, Application US/08959004
: Patent No. 6197543
: GENERAL INFORMATION:
: APPLICANT: Hillman, Jennifer L.
: APPLICANT: Yue, Henry
: APPLICANT: Corley, Neil C.
: APPLICANT: Lal, Preeti
: APPLICANT: Shah, Purvi
: APPLICANT: Kaser, Matthew
: TITLE OF INVENTION: HUMAN VESICLE MEMBRANE PROTEIN-LIKE
: TYPE OF INVENTION: PROTEINS
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Incyte Pharmaceuticals, Inc.
: STREET: 3174 Porter Drive
: CITY: Palo Alto
: STATE: CA
: COUNTRY: USA
: ZIP: 94304
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/959,004
: FILING DATE: Herewith
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Billings, Lucy J.
: REGISTRATION NUMBER: 36,749
: REFERENCE/DOCKET NUMBER: PF-0414 US
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650-855-0555
: TELEFAX: 650-845-4166
: TELEX:
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 667 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: LIBRARY: GenBank
: CLONE: 2131246
US-08-959-004-11

```

Query Match	22.5%	Score 694;	DB 4;	Length 667;
Best Local Similarity	27.0%;	Pred. No. 1.3e-53;		
Matches 175;	Conservative 116;	Mismatches 254;	Indels 102;	Gaps 14;
QY	21	TYQDKEEVLMNNTVGP-----YHNROE-----	---TYKFSLPFCVGVSKKSISHY	63
DB	32	TYRENDNIPLLVNLHTPSMNYQHKDEGNNVSGDKENFLYSYDYIYNNRPFECQEPKEVQ	91	
QY	64	HETLGEALQGVLEFSGLDIKFKDDVMPATVCEIDLQEKRDAPVYALKHNYWQMYIDD	123	
DB	92	PESIGSVIFGDRYIYNSPQLNMLQKECECSLCKTVIPGDDAKFINKLKNGFFQNWLLIDG	151	
QY	124	LP-----	---IWGITVGEADENGEDYVLTW---	146
DB	152	LPAAREVYDGRKTSFYGAGNGLGVQVTOGTDTIEATPKGAETTDKDOVELETRNDRNMVK	211	
QY	147	KLEIGFGNCR---IVDYNLTSSEGVKLV-----PNT-----	---	174
DB	212	TYELPYFANHFDMIEYHDRGEGRNYRVGVIVPSVKRISPSGTCETGSPMLMDEGNND	271	

```

QY 175 KIQMSYSVKKKSDVKFDRPKYL---DPSFTQHRTHWFSIPNSFMWVILFVLGYSMIL 231
DB 272 EIVETYSVKFNESATSWATRWKDKLHYVDPS-----IQWFSLINFSLVVLLSSVVIHSL 326
QY 232 MRTLKDYARYSKEEEDMDRLDGEYGNKQVHGDVFRSSPHLFSSLIGSGCOIFAV 291
DB 327 LRAUKSOFARYN-ELNLDL---DQEDSGWKLNHGDVFRSPQSLLTSLVGSVGLFLM 382
QY 292 SLIIVIAMIEDLY-TERGSMLSTAIYVAATSPVNGYFGGSLYARQGGRRNLIKQMFICA 350
DB 383 VTCSTFFAALGLFSPSSRGSLSATWFLYALFGVGSYTSWGIYKFNPGPYKKANLILTP 442
QY 351 FLIPAMVCGTRAFFINFNTAIYYHASRAIPFGTGMVAVCCICFFVILPNLVGTILGRNLSGO 410
DB 443 LLVFGAILLIITIALNFFLMFVSHGGVLPASTLFFMFLWFLFSLPSFAGSLIARKRCHW 502
QY 411 PNFCRVNAVPRPIPEKKWMEPAVIVCLGILLPFGSIFTEMYFIFTSWAKYIYVYVG 470
DB 503 DEHPKTNOIARQIPFPQWYKLTPTPATLIAGIIPFGSIAVELFYIYTSLWFKNIFMYFGF 562
QY 471 MMLVLVILCIVTCVTIVCYVYFLLNABDYRWQMTSF-LSAASTAIYVWYFYFYFFPKTK 529
DB 563 LFFSFLITLTSSIVTITLIYTHSLCLENWKNQWRGFTIGGACALYVFIHSI--LFTKFK 620
QY 530 MYGLFQTSFYFGYMAVFSTALGIMCGAIGYMGTSFAFRKIYTNVKKID 576
DB 621 LGGFTTIVLYGVSSVLSLLCCLVTGSGIGFISSMLFVRKIYSSIKVD 667

RESULT 4
US-09-134-001C-3703
; Sequence 3703, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO S
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3703
; LENGTH: 496
; TYPE: PRT
; ORGANISM: staphylococcus epidermidis
US-09-134-001C-3703

```

Query Match	3.8%	Score 118.5	DB 4	Length 496
Best Local Similarity	20.0%	Pred. No. 0.0035		
Matches 81	Conservative 68	Mismatches 156	Indels 99	Gaps
Qy 208	IHWFSIENSE--MMVIFVLGVLSMILMRLTKDYARYS--KKEEMDDMDRDLGDEYGNK 262	Db 7	MNYLKQVESFFWLIGIFLYLIMAILPLSTTDHAYKYNLSOYLQTQENGRYLGHLFEW 66	
Qy 263	QVHGCDVFRPSHPLITSSLLGSCQFAYS--LIVIVAMIEDLYTERGSMLSAIFVYA 320	Db 67	AVHNIIIR-----ALAYATISFLVLYVAYMQLHNFNREFYFILS--FVLM 109	
Qy 321	ATSPVNGYFGGSLYAROGGRWIKQWFIGAF-LIPAMVCGTAFFINFIAIYYHASRAIPE 379	Db 110	VTVP-----NTIYSEYTG--W--FTGPFSEYIPATV--LSLFLFTVVKWIESHD-- 152	
Qy 380	GTWVAVCCICFFVILPLNLVGTILGRNLGSPNPFQRCVNAVPRPPEKK--WFMEPAVI 436	Db 153	-----TVSEMGLWFLVLSLFGQPFLENLSANSLIILIGMVFYFVKRRLSYFLIVGFM 208	
Qy 437	VCLGGILPGCSIFTEMYFIF-----TSWAYKXYVYVGF--MLVL 475			

```
Db 209 SCIGNIMFLNF---NYFLIKDGLNTHYSISDSHGMIHKAGVTLFKLVPEYMFNQMIIL 265
QY 476 VILCIVTV-----CVYTI-VCTYFLLNAEDYRQWTSFLSAA- 510
Db 266 TVISIVSIVLLKONKSLKMRVYIKPLLGLLTPYIKFYVNOQHFELYKASFSIAVL 325
QY 511 -STAIYVYMSFYFFKTKMYGLFQTSFVGYMAVFSALGIM 553
Db 326 NTTICFIYMSIVVYVFKMQORVIRMVMSFIAMASSVLPLL 369

RESULT 5
US-09-134-001C-3057
; Sequence 3057, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3057
; LENGTH: 237
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3057
```

```
Query Match 3.6%; Score 112.5; DB 4; Length 237;
Best Local Similarity 20.7%; Pred. No. 0.0047;
Matches 51; Conservative 45; Mismatches 77; Indels 73; Gaps 12;

QY 352 LIPAMVGTAF-FT-----NFTAIYHASRAIPFGTMVAVCCIFPVILPLNLVGTILG 404
Db 3 LLLATYCSIIFLYQPCNQNIKLQVLYSR-----SHCDFCHTIIKPLDLLPIISF 54
QY 405 RNLSGQNPFCRNVAPRPIPEKKWFMEPAVIVCLGILPGFSIFTEMFYFTSFWAYKI 464
Db 55 LKLRGQSR--CCNQPLQR-----LVLIGELVSGAIFL-----YYPT 89
QY 465 YYVYGFNMLVLVICIVTVCTVYCYFLLNAEDYRQWNT-SFLSAASTAIY-----VY 517
Db 90 HFNHETFLTITFLTLTMCCLYDIHSMHI---DMRLLFYTVVSVFTTQTYGNETIMIF 145
QY 518 MYSFYFFKTKMYG-----LFQTSFVGYMAVFSALGIMCGAIGYMGTSAF 565
Db 146 LISHVYLFASKFIGYGDILLFNILGLIFLPLENFFF-FIVVFTFIIG-----GIFAI 195
QY 566 VRKIYT 571
Db 196 ILKIFS 201
```

```
RESULT 6
US-08-742-440A-6
; Sequence 6, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihari, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; NUMBER OF SEQUENCES: 3 and Uses Thereof
; CORRESPONDENCE ADDRESS: 23
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
```

```
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650-327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 408 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; US-08-742-440A-6

Query Match 3.5%; Score 107.5; DB 2; Length 408;
Best Local Similarity 18.9%; Pred. No. 0.032;
Matches 56; Conservative 43; Mismatches 122; Indels 75; Gaps 10;

QY 275 PLFSSLIIGSCQIFAVSLIVIIIVAMIEDLYTERGSMLSIAFVYAATSPVNGYFGSLY 334
Db 117 PAIYLLVYVGVANAVTLMLLFFRTSICITVYTNLAIADELFCVTLFP-----KIA 170
QY 335 ARGGRRWI-----KOMFTGAFILPAMVCGTAFINFIAYY-HASRAIPFGTMVA 384
Db 171 YHLNGNWNVEGELCRATTIVIFYGNMYCSILLACISINRYLAIVHPFTYRGLPKHTYAL 230
QY 385 VCC-----ICFFVLLPLNLVGTILGRNLSQPNPCPCRVNAVPRPIPEKKWFMEPAVIVC- 438
Db 231 VTCGLVNAVTVLYMLPFFI-----LKOEYLYVQPDITTC 265
QY 439 -LGGILPFGSIFTEMFYFTSFWAYKIYVYVGFMM-LVLVILCIVTVCTVICTYFLLNA 496
Db 266 DVHNTCESSPFQLYYFISLAF-----FGLIPVLIYCYAAIIRT-----LNA 310
QY 497 EYRWQW-----TSFLSAASTAIYVYMSFYFFKTKMYGLFQTSFYFG 541
Db 311 YDHRWLWYVKASLLILVIFTICEFAPSNIILIIHHANYVYNTDGLVFYIYIALCLG 366

RESULT 7
US-09-576-160B-6
; Sequence 6, Application US/09576160B
; Patent No. 6469150
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; APPLICANT: Zhao, Jiugao
; APPLICANT: Swanson, Robert
; APPLICANT: Webb, Maria
; APPLICANT: Strohl, Barbara
; TITLE OF INVENTION: Cloning and Characterization of Genes Encoding
; TITLE OF INVENTION: Bradykinin B1 Receptor Homologues From Five Mammalian
; TITLE OF INVENTION: Species
; FILE REFERENCE: 1073.058
; CURRENT APPLICATION NUMBER: US/09/576,160B
```

; REFERENCE/DOCKET NUMBER: 854-012 (32,141)

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440.674
; FILING DATE: May 15, 1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/107,347
; FILING DATE: August 16, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Alan M. Gordon
; REGISTRATION NUMBER: 30637
; REFERENCE/DOCKET NUMBER: 854-P0012 Div (32,141)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-831-3244
; TELEFAX: 201-831-3305
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: polypeptide
; FEATURE:
; NAME/KEY: Schizosaccharomyces pombe sts gene
; PUBLICATION INFORMATION:
; AUTHORS: M. Shimanuki, M. Goebel, M. Yanagida,
; AUTHORS: and T.
; AUTHORS: Toda
; TITLE: Fission Yeast stsl+ Gene
; TITLE: Encodes a
; TITLE: Protein Similar to the Chicken Lamin B Receptor
; JOURNAL: Molecu-
; JOURNAL: lar Biology of the Cell
; VOLUME: 3
; PAGES: 263-273
; PAGES: Sequence set out in Figure 1, page 264
; Patent No. 5525496
; DATE: 1992
; RELEVANT RESIDUES IN SEQ ID NO: open reading frame
; US-08-440-674-4

```

```

Query Match 3.4%; Score 104.5; DB 1; Length 453;
Best Local Similarity 20.3%; Pred. No. 0.074;
Matches 55; Conservative 46; Mismatches 103; Indels 67; Gaps 13;

QY 361 AFFINIAI-VYHSAIPAIP-----FGTMVAVCCICFFVILPLNLVGTIL-----G 404
Db 121 SFYTTIVLAVLVHVPITTFIDMFGLMSVAITAFVCTFVLTGLLFGDLFDKP 180
QY 405 RNLGQPNPCRVNAVPRP-----IPEKKWFMEPAV-----IVCLGGIL-----PFGSIFI 450
Db 181 HRLSGNPIDAFMGACLNPRLGLDKLFKMFVEVRIPWFLFTISVGAARQVETGTVSP 240
QY 451 ENYFIFTSWAY-----KIIYVYGFMMVLVLICVTVIVCVTVIV-CTYEL 493
Db 241 QVLFVCLGHLVYANACSGEQILIVTWDNAYEKFGFM-----LIFNMAGVPFTYSHCTIYL 297
QY 494 LNAEDYRWQTSFSLSAASTAIYVYVSYFYF-----FKTKMYG--LFQTSFYFYG 542
Db 298 FSHDPSVYVWS---TQYTTGIIVLLCCYYIFDTGCKNHFNRNIQYTEVHRKTFPQLP 354
QY 543 MAVFTALGIMCAIGYMGTSA---FVRKIY 570
Db 355 WLIIRKPTFIRCANGTLLTSGWRYARKIH 385

RESULT 10
PCT-US94-00198-3
; Sequence 3, Application PC/TUS9400198
; GENERAL INFORMATION:
; APPLICANT: Schering Corp.
; TITLE OF INVENTION: RAS Associated GAP Proteins
; NUMBER OF SEQUENCES: 6

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering Corp.
; STREET: 1 Girald Farms
; CITY: Madison
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: 6.0.8
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/00198
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/004,824
; FILING DATE: 15-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lunn, Paul G.
; REGISTRATION NUMBER: 32,743
; REFERENCE/DOCKET NUMBER: DX0352 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)822-7255
; TELEFAX: (201)822-7039
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2938 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Saccharomyces cerevisiae
; PCT-US94-00198-3

Query Match 3.3%; Score 101.5; DB 5; Length 2938;
Best Local Similarity 20.5%; Pred. No. 2.3;
Matches 99; Conservative 78; Mismatches 172; Indels 135; Gaps 26;

QY 132 EADENGEDYLLWTKKLEIGFNGNRIVDNLTSBGKVLKVPNTKIQMSYSVKWKSDV-- 189
Db 2466 ERNRKNDIVCACTFAKQKLN-----ISGFSQKGRVLPNFAAS-SFSSKFGTLDLFT 2517
QY 190 -----KFRDPRDKYLDPSFFQHRHWFISFNSEFMVIFVLGVSM----- 229
Db 2518 KNIMLLMEVGSISEGAQWEAKYKYLMDAIFGHR-----SFFSARAMMI--LGIMSKSHTS 2571
QY 230 -----ILMRTLKDYAR-----YSKEEEMDDMDRDLGDEYGNKQVHG 266
Db 2572 LFLCKELLVETM-KVFAEPVVDDEQMFIIAHVFTYSKIVEGLDPSSELKELFWL---A 2627
QY 267 DVFERSSHPLIFSSLLGSCQIFAVSLIIVIAMIEDLYTER-----CSMLSTAFIV-- 318
Db 2628 TTCVESPHPLLF-----EGLLLFMVN-----CLURLYTVHLQGLGFGDKSLAKLMESR 2675
QY 319 -YAAT--SPVNGYFGSGLYAROGRRWIKQMFIGAFLIP-----AMVCGTAFFINFIAY 370
Db 2676 NEAATLLAKLESY-NGCIWNEDNFPHIILGFTANGLSIPVVKGAALDCLQALFKN---T 2730
QY 371 YHASRAIPGTMVAVCCICFFVILPLNL-----VGTILGNLSQLQPNFCRVNAVPRPIPE 426
Db 2731 YVERKSNPKSSDY----LCYLEFLHLVLSPEQLSTLL-LEVGFEDLVPLNNTLVPLTL 2785
QY 427 KKWFF---MEPAVIVCLGGILPFGSIFIEFYFTSEFWAKIYVYVGFMMVLVLICIVTV 483
Db 2786 INWSSDSKSNIVLYQGALLFSCVSDSDPCKFR-----FALLMYLLKLVNPI 2833
QY 484 CVTIVCTYFLLNAEDYRWQTSFSLSAASTAIYVYVSYFYFFKTKMYGLFOTSFYFGYM 543
Db 2834 C---VFRFYTLTRKEER--RLSTLEQSSEAVAV---SF-----ELIGMLYTHSEFNVL 2878

```











GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 18, 2003, 20:52:42 ; Search time 40 seconds  
(without alignments)  
1918.809 Million cell updates/sec

Title: US-09-319-724A-14

Perfect score: 3089

Sequence: 1 AALMLLLLLLPRTRADEHH.....IGYMGTSFVRKIYTNVKID 576

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_101002.\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3089	100.0	579	21	Human SM-11044-bin
2	3089	100.0	582	21	Human SM-11044-bin
3	2923	94.6	545	22	Human protein sequ
4	2916	94.4	545	21	Human secreted pro
5	2786	90.2	530	22	Human secreted pro
6	2347	76.0	439	19	Non-adrenergic SM
7	2222.5	71.9	567	22	Drosophila melanog
8	2123	68.7	399	22	Human polypeptide,
9	1902	61.6	360	22	Human secreted pro
10	1757	56.9	329	22	Human polypeptide,

11	1757	56.9	329	22	AAB92687	Human protein sequ
12	1710	55.4	589	21	AAG21345	Arabidopsis thalia
13	1710	55.4	608	21	AAG21344	Arabidopsis thalia
14	1669	54.0	583	21	AAG45638	Arabidopsis thalia
15	1669	54.0	602	21	AAG45637	Arabidopsis thalia
16	1668	54.0	589	21	AAG30789	Arabidopsis thalia
17	1558	50.4	537	22	AAB75597	Gene 41 human seq
18	1519.5	49.2	496	21	AAG21346	Arabidopsis thalia
19	1478.5	47.9	490	21	AAG45639	Arabidopsis thalia
20	1477.5	47.8	496	21	AAG30790	Arabidopsis thalia
21	1426.5	46.2	461	21	AAG30791	Arabidopsis thalia
22	1148.5	37.2	592	21	AAG36371	Arabidopsis thalia
23	1146.5	37.1	592	21	AAG32077	Arabidopsis thalia
24	950.5	30.8	642	21	AA80519	Human putative SST
25	950.5	30.8	642	23	AAU74740	Human Phg1 protein
26	944	30.6	607	21	AAB38332	Human secreted pro
27	943.5	30.5	467	21	AAG36372	Arabidopsis thalia
28	938.5	30.4	467	21	AAG32078	Arabidopsis thalia
29	936.5	30.3	630	22	AAB63203	Drosophila melanog
30	898	29.1	609	21	AAG31959	Arabidopsis thalia
31	898	29.1	624	21	AAG31958	Arabidopsis thalia
32	898	29.1	659	21	AAG31957	Arabidopsis thalia
33	878	28.4	619	21	AAG43743	Arabidopsis thalia
34	860	27.8	641	23	AAU74739	Dictyostelium Phg1
35	859.5	27.8	416	21	AAG36373	Arabidopsis thalia
36	846.5	27.4	416	21	AAG32079	Arabidopsis thalia
37	843.5	27.3	663	20	AAV17390	Human vesicle memb
38	838	27.1	659	22	AAB58205	Drosophila melanog
39	812	26.3	218	23	AAB90443	Human polypeptide
40	810	26.2	517	21	AAG43745	Arabidopsis thalia
41	809	26.2	522	21	AAG43744	Arabidopsis thalia
42	789	25.5	509	21	AAG04118	Arabidopsis thalia
43	747	24.2	370	21	AAG04119	Arabidopsis thalia
44	715	23.1	152	22	AA93268	Human polypeptide,
45	700	22.7	356	21	AAG04120	Arabidopsis thalia

#### ALIGNMENTS

RESULT 1  
AA82460  
ID AA82460 standard; Protein; 579 AA.

XX AA82460;

XX AA82460;

DT 30-JUN-2000 (first entry)

XX Human SM-11044-binding receptor protein SEQ ID NO:4.

DE Human: SM-11044-binding receptor protein; SMAP; antiasthmatic;  
XX antiinflammatory; inflammation; eosinophil infiltration; asthma;  
KW intestinal disease; eosinophil migration inhibitor; intestinal stress.

XX Homo sapiens.

XX OS

XX WO200014266-A1.

PN 16-MAR-2000.

PD 06-SEP-1999; 99WO-JP04808.

XX 08-SEP-1998; 98JP-0253771.

PR (SUMU ) SUMITOMO PHARM CO LTD.

XX Sugawara T, Hidaka J, Kawakami H;

XX WPI: 2000-257008/22.

DR N-PSDB; AAA08361.

XX Recombinant human SM-11044-binding receptor protein with ligand binding  
PT activity, useful for developing drugs for inflammation accompanying

PT eosinophil infiltration, asthma and intestinal diseases -  
XX  
PS  
XX Example 3; Page 36-39; 41pp; Japanese.

CC The present invention describes transformed cells which express a  
CC recombinant human SM-11044-binding receptor protein (SMBP) at such  
CC a high level as to enable the assay of the ligand-binding activity by  
CC deleting the polythymidine sequence from the base sequence of the  
CC 3'-nontranslation region or the cell membrane fraction. Also described  
CC is a method for screening a human SMBP agonist/antagonist characterised  
CC by using the transformed cells, cell membrane fraction of the  
CC recombinant human SMBP. The recombinant protein is useful for developing  
CC drugs for treating inflammation accompanying eosinophil infiltration,  
CC asthma, and diseases of intestine, particularly agonists as eosinophil  
CC migration inhibitors or relaxing agents for intestinal stress. The  
CC protein is easily expressed by transformed cells after deleting  
CC polythymidine sequence from base sequence of 3'-nontranslation region.  
CC The present sequence represents human SMBP from the present invention.  
XX  
SQ Sequence 579 AA;

Query Match 100.0%; Score 3089; DB 21; Length 579;  
Best Local Similarity 100.0%; Pred. No. 1.9e-301;  
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALWLLLLLPRTRADEHEHTYQDKKEEVLWMTVGPYHNROETKYKYSLPFCVGSKKSI 60  
DB 4 AALWLLLLLPRTRADEHEHTYQDKKEEVLWMTVGPYHNROETKYKYSLPFCVGSKKSI 63

QY 61 SHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLKDKRDFAFYAIAKNHYQMY 120  
DB 64 SHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLKDKRDFAFYAIAKNHYQMY 123

QY 121 IDDLPTWIGVEADENGEDYLLWTYKKLEIGFNGNRIVDVNLTSSEKVKLVNPTKQMSY 180  
DB 124 IDDLPTWIGVEADENGEDYLLWTYKKLEIGFNGNRIVDVNLTSSEKVKLVNPTKQMSY 183

QY 181 SVKWKSDVKFEDRDKYLDPSFQHRHWFSTFNSFMVIFLVGLVSMILMRTLKRDYA 240  
DB 184 SVKWKSDVKFEDRDKYLDPSFQHRHWFSTFNSFMVIFLVGLVSMILMRTLKRDYA 243

QY 241 RYSKEEMDDMDRLDGEYQVHGDVFRPSSHPLIFSSLGSGCQIFAVSLIIVIAM 300  
DB 244 RYSKEEMDDMDRLDGEYQVHGDVFRPSSHPLIFSSLGSGCQIFAVSLIIVIAM 303

QY 301 IEDLYTERGSMLSLTAIFYAATSPVNGYFGGSLYARQGGRRWIKQMFIAFLIPAMVCGT 360  
DB 304 IEDLYTERGSMLSLTAIFYAATSPVNGYFGGSLYARQGGRRWIKQMFIAFLIPAMVCGT 363

QY 361 AFFINFIAIYYHASRAIPFGTAVVAVCCICFFVLPLNLVGTILGRNLGQPNPFCRVNAV 420  
DB 364 AFFINFIAIYYHASRAIPFGTAVVAVCCICFFVLPLNLVGTILGRNLGQPNPFCRVNAV 423

QY 421 PRPIPEKKWFMPEAVTVCLGGILPFGSFIEMFYFTSFWAYKIYVYVGFMMVLVLVLCI 480  
DB 424 PRPIPEKKWFMPEAVTVCLGGILPFGSFIEMFYFTSFWAYKIYVYVGFMMVLVLVLCI 483

QY 481 VTCVTVICTYFLNADRYWQNTSFLSAATAIYVYVYVSEFYFEKTKWYGLQFSFYF 540  
DB 484 VTCVTVICTYFLNADRYWQNTSFLSAATAIYVYVYVSEFYFEKTKWYGLQFSFYF 543

QY 541 GYMAVFESTALGIMCGAIGYMGTSFAVRKIYTVNKKID 576  
DB 544 GYMAVFESTALGIMCGAIGYMGTSFAVRKIYTVNKKID 579

RESULT 2  
ID AAY82459  
XX AAY82459 standard; Protein; 582 AA.  
AC AAY82459;  
XX  
DT 30-JUN-2000 (first entry)

XX Human SM-11044-binding receptor protein SEQ ID NO:2.  
DE  
XX  
KW Human; SM-11044-binding receptor protein; SMBP; antiasthmatic;  
KW antiinflammatory; inflammation; eosinophil infiltration; asthma;  
XX intestinal disease; eosinophil migration inhibitor; intestinal stress.  
OS  
XX Homo sapiens.  
PN WO200014266-A1.  
XX 16-MAR-2000.  
PD  
XX 06-SEP-1999; 99WO-JP04808.  
PF  
XX 08-SEP-1998; 98JP-0253771.  
PR  
XX (SUMU ) SUMITOMO PHARM CO LTD.  
PA  
XX Sugasawa T, Hidaka J, Kawakami H;  
PI  
XX WPI: 2000-257008/22.  
DR N-PSDB; AAA08360.  
XX  
PT Recombinant human SM-11044-binding receptor protein with ligand binding  
PT activity, useful for developing drugs for inflammation accompanying  
PT eosinophil infiltration, asthma and intestinal diseases -  
XX  
PS Example 1; Page 29-32; 41pp; Japanese.

CC The present invention describes transformed cells which express a  
CC recombinant human SM-11044-binding receptor protein (SMBP) at such  
CC a high level as to enable the assay of the ligand-binding activity by  
CC deleting the polythymidine sequence from the base sequence of the  
CC 3'-nontranslation region or the cell membrane fraction. Also described  
CC is a method for screening a human SMBP agonist/antagonist characterised  
CC by using the transformed cells, cell membrane fraction of the  
CC recombinant human SMBP. The recombinant protein is useful for developing  
CC drugs for treating inflammation accompanying eosinophil infiltration,  
CC asthma, and diseases of intestine, particularly agonists as eosinophil  
CC migration inhibitors or relaxing agents for intestinal stress. The  
CC protein is easily expressed by transformed cells after deleting  
CC polythymidine sequence from base sequence of 3'-nontranslation region.  
CC The present sequence represents human SMBP from the present invention.  
XX  
SQ Sequence 582 AA;

Query Match 100.0%; Score 3089; DB 21; Length 582;  
Best Local Similarity 100.0%; Pred. No. 1.9e-301;  
Matches 576; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AALWLLLLLPRTRADEHEHTYQDKKEEVLWMTVGPYHNROETKYKYSLPFCVGSKKSI 60  
DB 7 AALWLLLLLPRTRADEHEHTYQDKKEEVLWMTVGPYHNROETKYKYSLPFCVGSKKSI 66

QY 61 SHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLKDKRDFAFYAIAKNHYQMY 120  
DB 67 SHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLKDKRDFAFYAIAKNHYQMY 126

QY 121 IDDLPTWIGVEADENGEDYLLWTYKKLEIGFNGNRIVDVNLTSSEKVKLVNPTKQMSY 180  
DB 127 IDDLPTWIGVEADENGEDYLLWTYKKLEIGFNGNRIVDVNLTSSEKVKLVNPTKQMSY 186

QY 181 SVKWKSDVKFEDRDKYLDPSFQHRHWFSTFNSFMVIFLVGLVSMILMRTLKRDYA 240  
DB 187 SVKWKSDVKFEDRDKYLDPSFQHRHWFSTFNSFMVIFLVGLVSMILMRTLKRDYA 246

QY 241 RYSKEEMDDMDRLDGEYQVHGDVFRPSSHPLIFSSLGSGCQIFAVSLIIVIAM 300  
DB 247 RYSKEEMDDMDRLDGEYQVHGDVFRPSSHPLIFSSLGSGCQIFAVSLIIVIAM 306

QY 301 IEDLYTERGSMLSLTAIFYAATSPVNGYFGGSLYARQGGRRWIKQMFIAFLIPAMVCGT 360

Db 307 IEDLYTERGSMSTAIFFVYAANTSPVNGYFGGSLYARQGGRRWIKQMFIGAFLIPAMVCGT 366

QY 361 AFFINFIAIYYHASRAIPFGTMAVCCICFFVILPLNLVGTILGRNLSCGNPFCRVNAV 420

Db 367 AFFINFIAIYYHASRAIPFGTMAVCCICFFVILPLNLVGTILGRNLSCGNPFCRVNAV 426

QY 421 PRPIPEKKWMEPAVIVCLGGILPFGSIFTEMVFIFTSFWAYKIYYVYGFMMVLVLICI 480

Db 427 PRPIPEKKWMEPAVIVCLGGILPFGSIFTEMVFIFTSFWAYKIYYVYGFMMVLVLICI 486

QY 481 VTVCVTIVCTYFLNADRYRWQNTSFLSAASTAIYVYMYSFYFFPKTKMYGLFQTSFYF 540

Db 487 VTVCVTIVCTYFLNADRYRWQNTSFLSAASTAIYVYMYSFYFFPKTKMYGLFQTSFYF 546

QY 541 GYMAVFSTALGIMCGAIGYMGTSFAVRKIYTNVKID 576

Db 547 GYMAVFSTALGIMCGAIGYMGTSFAVRKIYTNVKID 582

RESULT 3

ID AAB94667 standard; Protein; 545 AA.

XX AAB94667;

XX

DT 26-JUN-2001 (first entry)

XX

DE Human protein sequence SEQ ID NO:15600.

XX

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX

OS Homo sapiens.

XX

PN EP1074617-A2.

XX

PD 07-FEB-2001.

XX

PF 28-JUL-2000; 2000EP-0116126.

XX

PR 29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.

PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.

PR 09-JUN-2000; 2000JP-0241899.

XX

PA (HELI-) HELIX RES INST.

XX

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX

DR WPI; 2001-318749/34.

XX

PT Primer sets for synthesizing polynucleotides, particularly the 5602

PT full-length cDNAs defined in the specification, and for the detection

PT and/or diagnosis of the abnormality of the proteins encoded by the

XX full-length cDNAs -

PS Claim 8; SEQ ID 15600; 2537pp + CD ROM; English.

XX

CC The present invention describes primer sets for synthesising 5602

CC full-length cDNAs defined in the specification. Where a primer set

CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary

CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the

CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination

CC of an oligonucleotide comprising a sequence complementary to the

CC complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a

CC polynucleotide which comprises a 3'-end sequence, where the

CC oligonucleotide comprises at least 15 nucleotides and the combination of

CC the 5'-end sequence/3'-end sequence is selected from those defined in

CC the specification. The primer sets can be used in antisense therapy and

CC in gene therapy. The primers are useful for synthesising polynucleotides,

CC particularly full-length cDNAs. The primers are also useful for the

CC detection and/or diagnosis of the abnormality of the proteins encoded by;

CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and

CC AAH13633 to AAH18742 represent human cDNA sequences; AAB9446 to

CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632

CC represent oligonucleotides, all of which are used in the exemplification;

XX of the present invention.

QY Sequence 545 AA;

Query Match 94.6%; Score 2923; DB 22; Length 545;

Best Local Similarity 100.0%; Pred. No. 8.5e-285;

Matches 545; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 MNTVGPYHNRQETKYKFSLPFCVSGSKKSIHSHYHETLGEALOGVELEFSGLDIKFKDDVMP 91

Db 1 MNTVGPYHNRQETKYKFSLPFCVSGSKKSIHSHYHETLGEALOGVELEFSGLDIKFKDDVMP 60

QY 92 ATYCEIDLQEKRDADFVYAIKKNHYWYQMYIDDLPIWGIWGEADENGEDYLYWTKKLEIG 151

Db 61 ATYCEIDLQEKRDADFVYAIKKNHYWYQMYIDDLPIWGIWGEADENGEDYLYWTKKLEIG 120

QY 152 FNGNRIVDNLTSSEGVKVLVPNTKIOMSYSVKWKSDVKFEDRFDKYLDPSPFQHRHWF 211

Db 121 FNGNRIVDNLTSSEGVKVLVPNTKIOMSYSVKWKSDVKFEDRFDKYLDPSPFQHRHWF 180

QY 212 SFNSPMWVIFLVGLVSMILMRTLKDYARYSKEEEMDDMDRLDGEYGWKQVHGDFRP 271

Db 181 SFNSPMWVIFLVGLVSMILMRTLKDYARYSKEEEMDDMDRLDGEYGWKQVHGDFRP 240

QY 272 SSHPLIFSSLSIGGCOIFAVSLIIVIAMIEDLYTERGSMSTAIFFVYAANTSPVNGYFGG 331

Db 241 SSHPLIFSSLSIGGCOIFAVSLIIVIAMIEDLYTERGSMSTAIFFVYAANTSPVNGYFGG 300

QY 332 SLYARQGGRRWIKQMFIGAFLIPAMVCGTAFFINFIAIYYHASRAIPFGTMAVCCICFF 391

Db 301 SLYARQGGRRWIKQMFIGAFLIPAMVCGTAFFINFIAIYYHASRAIPFGTMAVCCICFF 360

QY 392 VILPLNLVGTILGRNLSCGNPFCRVNAVPRPIPEKKWMEPAVIVCLGGILPFGSIFTE 451

Db 361 VILPLNLVGTILGRNLSCGNPFCRVNAVPRPIPEKKWMEPAVIVCLGGILPFGSIFTE 420

QY 452 MYFIFTSFWAYKIYYVYGFMMVLVLICITVCTVTCVTFLLNADRYRWQNTSFLSAAS 511

Db 421 MYFIFTSFWAYKIYYVYGFMMVLVLICITVCTVTCVTFLLNADRYRWQNTSFLSAAS 480

QY 512 TAIYVYMYSFYFFPKTKMYGLFQTSFYFGYMAVFSTALGIMCGAIGYMGTSFAVRKIY 571

Db 481 TAIYVYMYSFYFFPKTKMYGLFQTSFYFGYMAVFSTALGIMCGAIGYMGTSFAVRKIY 540

QY 572 NVKID 576

Db 541 NVKID 545

RESULT 4

ID AAY94910 standard; Protein; 545 AA.

XX AAY94910;

XX

DT 16-JUN-2000 (first entry)

XX

DE Human secreted protein clone pk366\_7 protein sequence SEQ ID NO:26.

XX

KW Human; secreted protein; immunestimulant; immunosuppressant; virucide;

KW antibacterial; antifungal; cytostatic; antiinflammatory; dermatological;

KW antidiabetic; antiasthmatic; antiarthritic; antirheumatic; protozoacide;

KW antihistoid; immune deficiency; severe combined immunodeficiency; SCID;

KW infection; HIV; hepatitis; malaria; autoimmune disorder; systemic lupus;

KW connective tissue disease; multiple sclerosis; erythematosis;

KW rheumatoid arthritis; autoimmune pulmonary inflammation; asthma;

KW Guillain-Barre syndrome; autoimmune thyroiditis; myasthenia gravis;  
KW insulin dependent diabetes mellitus; graft-versus-host-disease;  
KW autoimmune inflammatory eye disease; allergy.  
OS Homo sapiens.  
XX WO200009552-A1.  
PN  
PD 24-FEB-2000.  
XX  
XX 13-AUG-1999; 99WO-US18298.  
XX  
PR 14-AUG-1998; 98US-0096622.  
PR 17-AUG-1998; 98US-0096815.  
PR 04-SEP-1998; 98US-0099229.  
PR 23-OCT-1998; 98US-0105368.  
PR 08-JAN-1999; 99US-0115234.  
PR 12-FEB-1999; 99US-0119931.  
PR 18-FEB-1999; 99US-0120575.  
PR 30-APR-1999; 99US-0132020.  
PR 11-AUG-1999; 99US-0096622.  
XX (GEMY ) GENETICS INST INC.  
PA  
XX Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Agostino MJ, Steininger RJ, Spaulding V;  
PI Wong GG, Clark HF, Fechtel K;  
XX WPI: 2000-205979/18.  
XX  
XX New polynucleotides encoding secreted proteins, which may have e.g.  
PT nutritional, chemokine, immune stimulating or suppressing,  
PT hematopoiesis regulating, tissue growth, activin/inhibin  
PT antiinflammatory or tumor inhibition activity  
XX  
PS Claim 35; Page 494-495; 641pp; English.  
XX  
XX AAA16618 to AAA16697 encode the human secreted proteins given in  
CC AAY9498 to AAY9498, isolated from human adult brain, adult thyroid,  
CC adult retina, foetal carcinoma, adult blood, adult neural, foetal kidney,  
CC adult placenta, adult testis, whole embryo, adult cartilage, kidney,  
CC foetal brain, adult thymus, foetal placenta, adult uterus, adult tumour,  
CC and adult bladder, cDNA libraries. The polynucleotides and proteins are  
CC predicted to have biological activities which would make them suitable  
CC for treating, preventing or ameliorating medical conditions in humans  
CC and animals. The polynucleotides can be used as markers for tissues in  
CC which the protein is preferentially expressed, as molecular weight  
CC markers on Southern gels, and as chromosome markers or tags to identify  
CC chromosomes or to map gene positions. The proteins can be used in the  
CC treatment of immune deficiencies and disorders, such as severe combined  
CC immunodeficiency (SCID), as well as viral, bacterial, fungal and other  
CC infections. These infections include human immunodeficiency virus (HIV),  
CC hepatitis, herpesviruses, mycobacteria, Leishmania spp., malaria and  
CC candidiasis. The proteins can be used to treat autoimmune disorders such  
CC as connective tissue disease, multiple sclerosis, systemic lupus  
CC erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation,  
CC Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent  
CC diabetes mellitus, myasthenia gravis, graft-versus-host-disease and  
CC autoimmune inflammatory eye disease. The proteins can also be used to  
CC treat allergic conditions, such as asthma. AAA16698 to AAA16774 represent  
CC probes for the human secreted proteins from the present invention.  
XX  
SQ Sequence 545 AA:  
Query Match 94.4%; Score 2916; DB 21; Length 545;  
Best Local Similarity 99.8%; Pred. No. 4.3e-284;  
Matches 544; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 32 MNTVGPYHNROETKYFSLPCVCGSKSISHYHETLGEALOGVELEFSGLDIKFKDDVMP 91  
DB 1 MNTVGPYHNROETKYFSLPCVCGSKSISHYHETLGEALOGVELEFSGLDIKFKDDVMP 60  
OY 92 ATYCEIDLKDKRDFAFVAIKNHVYQWYIDDLPIWGIAGEADENGEDYLLWYTKKLEIG 151

Db 61 ATYCEIDLKDKRDFAFVAIKNHVYQWYIDDLPIWGIAGEADENGEDYLLWYTKKLEIG 120  
OY 152 ENGRIVDVNLTSSEGVKLVNPTKIOMSYSVKWKSKDKFEDRFDKYLDPSPFQRIHWF 211  
Db 121 FNGNRIVDVNLTSSEGVKLVNPTKIOMSYSVKWKSKDKFEDRFDKYLDPSPFQRIHWF 180  
OY 212 SFNSFMVIFLVGLVSMILMRTLRKDYARYSKEREMDDMDRLDGLDEYGWQVHGDFRP 271  
Db 181 SFNSFMVIFLVGLVSMILMRTLRKDYARYSKEREMDDMDRLDGLDEYGWQVHGDFRP 240  
OY 272 SSHPLIFSSLGSCQIFAVSLIIVIAMIEDLYTERGSMLSLTAIFVYAATSPVNGYFEG 331  
Db 241 SSHPLIFSSLGSCQIFAVSLIIVIAMIEDLYTERGSMLSLTAIFVYAATSPVNGYFEG 300  
OY 332 SLIARQGGRRWIKOMFIFAGFLIPAMVCGTAFFINIAIYYHASRAIPFGTMMVAVCCICFF 391  
Db 301 SLIARQGGRRWIKOMFIFAGFLIPAMVCGTAFFINIAIYYHASRAIPFGTMMVAVCCICFF 360  
OY 392 VILPLNLVGTILGRNLGQPNFPCRVNAVPRPIPEKKWFMPEPAVIVCLGGILPFGSIFIE 451  
Db 361 VILPLNLVGTILGRNLGQPNFPCRVNAVPRPIPEKKWFMPEPAVIVCLGGILPFGSIFIE 420  
OY 452 MYFTPTSWAKIYYVYGFMMVLVLICIVTVCVTVICTVYFLLNAEDYRWQNTSFLSAAS 511  
Db 421 MYFTPTSWAKIYYVYGFMMVLVLICIVTVCVTVICTVYFLLNAEDYRWQNTSFLSAAS 480  
OY 512 TAIYVYMYSFYFFPKTKMYGLFQTSFYFGYMAVFEVSTALGMCAGYMGTSFAVRKIYT 571  
Db 481 TAIYVYMYSFYFFPKTKMYGLFQTSFYFGYMAVFEVSTALGMCAGYMGTSFAVRKIYT 540  
OY 572 NVKID 576  
Db 541 NVKID 545  
RESULT 5  
ID AAB75598 standard; Protein; 530 AA.  
XX AAB75598;  
AC AAB75598;  
XX  
DT 06-APR-2001 (first entry)  
XX  
DE Human secreted protein sequence encoded by gene 41 SEQ ID NO:152.  
XX  
KW Human; secreted protein; immunosuppressive; antiarthritic; anti-rheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neutrophic; neuroprotective; antibacterial; virucide; fungicide;  
KW ophthalmological; vulnary; autoimmune disease; cardiovascular disorder;  
KW hyperproliferative disorders; cerebrovascular disorder; wound healing;  
KW nervous system disorder; ocular disorder; skin ageing; chemotaxis;  
KW food additive.  
XX  
XX Homo sapiens.  
OS  
XX WO200077026-A1.  
PN  
PD 21-DEC-2000.  
XX  
XX 01-JUN-2000; 2000WO-US14973.  
PF  
XX 11-JUN-1999; 99US-0138630.  
PR  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA (ROSE/) ROSEN C A.  
XX  
XX Rosen CA, Ruben SM, Komatsoulis GA;  
PI  
XX WPI: 2001-071258/08.  
DR N-PSDB; AAF64216.  
XX  
XX Nucleic acid molecules encoding human secreted proteins, used in



preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers -

Disclosure; Page 85-86; 542pp; English.

Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide sequences AAF64176 - AAF64224. The specification includes amino acid sequences AAB75555 - AAB75606 which represent fragments of the human secreted proteins, and protein sequences with which they share homology. The proteins and polynucleotides, their agonists and antagonists have activities dependent on the tissues and cells in which they are expressed, examples of these activities include, immunosuppressive; antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective; antibacterial; virucide; fungicide; opthalmological; and vulnery. The proteins, polynucleotides, agonists and antagonists can be used to treat or detect or diagnose various diseases and disorders including, autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease, infections caused by bacteria, viruses and fungi and ocular disorders e.g. corneal infection. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin ageing due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities. Included in the invention are polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which are used in the isolation, identification and characterisation of the proteins of the invention.

Sequence 530 AA;

Query Match	90.2%;	Score 2786;	DB 22;	Length 530;
Best Local Similarity	99.8%;	Pred. No. 4.9e-271;		
Matches 521;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
XX	55	GSKKSISHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLDKRDFAFYATKNH	114	
DB	9	GQKKSISHYHETLGEALQGVLEFSGLDIKFKDDVMPATYCEIDLDKRDFAFYATKNH	68	
QY	115	YWYOMYIDDLPIWGIVGEADENGEDYYLWYKKLEIGFNGNRIVDVNLTSEGVKVLVPNT	174	
DB	69	YWYOMYIDDLPIWGIVGEADENGEDYYLWYKKLEIGFNGNRIVDVNLTSEGVKVLVPNT	128	
QY	175	KIOMSYSVKWKSDVKFEDRDKYLDPSFTQHRHWFSEIFNSFMVIFLVGLVSMILMRT	234	
DB	129	KIOMSYSVKWKSDVKFEDRDKYLDPSFTQHRHWFSEIFNSFMVIFLVGLVSMILMRT	188	
QY	235	LRKDYARYSKEEEMDDMDRLDGEYGNKQVHGDVFRPSSHPLIFSSLLIGSCQIFAVSLI	294	
DB	189	LRKDYARYSKEEEMDDMDRLDGEYGNKQVHGDVFRPSSHPLIFSSLLIGSCQIFAVSLI	248	
QY	295	VIIVAMIEDLYTERGSMSTAIFYAATSPVNGYFGGSLYARQGRRWIKOMFIGAFLIP	354	
DB	249	VIIVAMIEDLYTERGSMSTAIFYAATSPVNGYFGGSLYARQGRRWIKOMFIGAFLIP	308	
QY	355	AMVCGTAFFNFNTAIYYHASRAIPFGTMVAVCCICFFVILPLNLVGTILGRNLSGQNF	414	
DB	309	AMVCGTAFFNFNTAIYYHASRAIPFGTMVAVCCICFFVILPLNLVGTILGRNLSGQNF	368	
QY	415	CRVNAVPRPIPEKKWFPEPAVIVCLGILPFGSIFTMYEFTSFWAYKIYYVYGFMMVL	474	
DB	369	CRVNAVPRPIPEKKWFPEPAVIVCLGILPFGSIFTMYEFTSFWAYKIYYVYGFMMVL	428	
QY	475	LVILCIVTCVTIVCTVYFLLNAEDYRWQWTSFSLAASATAIYVYMYSFYFFFKTKMYGLF	534	
DB	429	LVILCIVTCVTIVCTVYFLLNAEDYRWQWTSFSLAASATAIYVYMYSFYFFFKTKMYGLF	488	
QY	535	QTSFYFGYNAVFTALGIMCGAICYMGTSFAVRKIYTNVKID	576	
DB	489	QTSFYFGYNAVFTALGIMCGAICYMGTSFAVRKIYTNVKID	530	

RESULT 6

AAW61371 standard; protein; 439 AA.

XX AAW61371;

DT 25-SEP-1998 (first entry)

XX Non-adrenergic SM binding protein.

XX Non-adrenergic SM binding protein; human; muscle; receptor; idocyanopindolol.

OS Homo sapiens.

PN EP848059-A1.

XX 17-JUN-1998.

XX 12-DEC-1996; 96EP-0402719.

XX 12-DEC-1996; 96EP-0402719.

PA (VETI-) VETIGEN.

PI Lenzen G, Morooka S, Strosberg AD, Sugawara T;

DR WPI; 1998-314474/28.

DR N-PSDB; AAV28120.

XX New idocyanopindolol receptor polypeptide - and corresponding DNA, antibodies, etc.

PS Claim 2; Page 20-22; 56pp; English.

XX The non-adrenergic SM binding protein is found in human muscles. It is a member of a novel receptor class which have a receptor activity other than that of beta<sub>2</sub>-adrenergic receptors. The protein comprises sites which, when exposed on the cell surface, are capable of binding idocyanopindolol (ICYP) under blockade of alpha<sub>1</sub>, beta<sub>1</sub>, beta<sub>2</sub> and beta<sub>3</sub> adrenergic receptors and serotonin 5-HT<sub>1A</sub> and 5-HT<sub>1B</sub> receptors. The binding is saturable and reversible by displacement with the stereoselective beta-adrenergic receptor agonist SM-11044 but not isoproterenol, epinephrine, norepinephrine, serotonin, dopamine or BRL-37344.

XX Sequence 439 AA;

Query Match 76.0%; Score 2347; DB 19; Length 439;  
Best Local Similarity 100.0%; Pred. No. 5.1e-227;  
Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 MYIDDLPIWGIVGEADENGEDYYLWYKKLEIGFNGNRIVDVNLTSEGVKVLVPNTKIOM 178

DB 1 MYIDDLPIWGIVGEADENGEDYYLWYKKLEIGFNGNRIVDVNLTSEGVKVLVPNTKIOM 60

QY 179 SYSVKWKSDVKFEDRDKYLDPSFTQHRHWFSEIFNSFMVIFLVGLVSMILMRTLRKD 238

DB 61 SYSVKWKSDVKFEDRDKYLDPSFTQHRHWFSEIFNSFMVIFLVGLVSMILMRTLRKD 120

QY 239 YARYSKEEEMDDMDRLDGEYGNKQVHGDVFRPSSHPLIFSSLLIGSCQIFAVSLIIV 298

DB 121 YARYSKEEEMDDMDRLDGEYGNKQVHGDVFRPSSHPLIFSSLLIGSCQIFAVSLIIV 180

QY 299 AMIEDLYTERGSMSTAIFYAATSPVNGYFGGSLYARQGRRWIKOMFIGAFLIPAMVC 358

DB 181 AMIEDLYTERGSMSTAIFYAATSPVNGYFGGSLYARQGRRWIKOMFIGAFLIPAMVC 240

QY 359 GTAFFINFIAIYYHASRAIPFGTMVAVCCICFFVILPLNLVGTILGRNLSGQNFPCRVN 418

DB 241 GTAFFINFIAIYYHASRAIPFGTMVAVCCICFFVILPLNLVGTILGRNLSGQNFPCRVN 300

QY 419 AVPRPIPEKKWMEPAVIVCLGILPFGSIFIDMYFIETSFNAYKIYVYVGFMMVLVL 478  
DB 301 AVPRPIPEKKWMEPAVIVCLGILPFGSIFIDMYFIETSFNAYKIYVYVGFMMVLVL 360  
QY 479 CIVTVCVTIVCTYFLNAEDYRNQWTSFLSAASTAIYVYMYSFYFFKTKMYGLQTSF 538  
DB 361 CIVTVCVTIVCTYFLNAEDYRNQWTSFLSAASTAIYVYMYSFYFFKTKMYGLQTSF 420  
QY 539 YFGYMAVESTALGIMGAI 557  
DB 421 YFGYMAVESTALGIMGAI 439  
RESULT 7  
ID ABB65162 standard; Protein; 567 AA.  
XX  
AC ABB65162;  
DT 26-MAR-2002 (first entry)  
DE Drosophila melanogaster polypeptide SEQ ID NO 22278.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR N-PSDB; ABL09265.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 22278; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (AB857737-AB872072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 567 AA;  
Query Match 71.9%; Score 2222.5; DB 22; Length 567;  
Best Local Similarity 70.1%; Pred. No. 2.4e-214;  
Matches 406; Conservative 64; Mismatches 80; Indels 29; Gaps 2;  
QY 2 ALWLLLLL---LPRTRDEHEITYODKEEVLMWMTVGPYHNRQETKYFSLPFCVGSK 57  
DB 14 ATCLCLLIASCYVLSQADEHNKYNDRREVLMWMTVGPYHNRQETKYFSLPFCVGSK 73

QY 58 KSISHVHETLGEALQVELEFSGLDIKFKDDVNPATPYCEIDLDKKRDAFVAIAKNHYWY 117  
DB 74 SSISHVHETLSEALQVELEFSGYEMEFKSDAPKSKIVMTLQESAKATVAKNEYWY 133  
QY 118 QMYIDPLPWIGVGEADENGEDYLLWYKLEIGFNGNRIVDNLVNSEGKVLVPNTKIQ 177  
DB 134 QMYIDGLPWGKVGDERDGRKYIIFTHKKFDIGYNGQQIVDITLTEGREELKPGSHIN 193  
QY 178 MSYSVKKWKSDFKDFDKYLDPSFFQRIHWFSEFNSFMVIFLVGLVSMILMTLRK 237  
DB 194 FSYEVNWKPSKVEFKNRDKYLDPNFQRIHWFSEFNSFMVIFLVGLVSMILMTLRK 253  
QY 238 DYARYSKEEEMDDMDRLDGEYGWKQVHGVDPSPSSPHLIFSSLGSGCQIFAVSLIVII 297  
DB 254 DYARYSKEEIDDMERDLGDEYGWKQVHGVDPSPPTLLFSALVGAGYOLISVVFCEIM 313  
QY 298 VAMIEDLYTERGSMSTAIIFYAATSPVANGYFGSLYARQGGRRWIKOMFIGNFLIPAMV 357  
DB 314 FAIVGELYTERGSMSTAIIFYAATSPINGYFGSLYARLGGRRWIRQMLVSAFTYPAV 373  
QY 358 CGTAFPIFIATYVHASRAIPGTMVAVCCICFFVILPLNLVGTILGRNLGQPNPCRV 417  
DB 374 CGTAFILNFIATYVHASRAIPGTMVAVTCICLVILPLTVGTV----- 418  
QY 418 NAVPRPIPEKKWMEPAVIVCLGILPFGSIFIDMYFIETSFNAYKIYVYVGFMMVLVL 477  
DB 419 -----KWYMEPLIIVLLGGVLPFGSIFIDMYFIETSFNAYKIYVYVGFMMVLVFSI 468  
QY 478 LCIVIVCVTIVCTYFLNAEDYRNQWTSFLSAASTAIYVYMYSFYFFKTKMYGLQFS 537  
DB 469 LTVVTVCVTIVCTYFLNAEDYRNQWTSFMAAGSTSIYVYAYSFYFFKTKMFGLEFOTA 528  
QY 538 FYFGYMAVESTALGIMGAIYGMGTSAFYRKIVTNVKID 576  
DB 529 FYFGYMAVESTALGIMGAIYGMGTSAFYRKIVTNVKID 567  
RESULT 8  
ID AAM93355 standard; Protein; 399 AA.  
XX  
AC AAM93355;  
DT 06-NOV-2001 (first entry)  
DE Human polypeptide, SEQ ID NO: 2909.  
XX  
KW Human; full length cDNA; cDNA synthesis; oligo-capping.  
XX  
OS Homo sapiens.  
XX  
PN EP1130094-A2.  
XX  
PD 05-SEP-2001.  
XX  
PF 07-JUL-2000; 2000EP-0114089.  
XX  
PR 08-JUL-1999; 99JP-0194486.  
PR 11-JAN-2000; 2000JP-0118774.  
PR 02-MAY-2000; 2000JP-0183765.  
XX  
PA (HELI-) HELIX RES INST.  
XX  
PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
DR WPI; 2001-524255/58.  
DR N-PSDB; AAK94275.  
XX  
PT 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
XX  
PS Claim 8; SEQ ID NO 2909; 1380pp + sequence listing; English.

XX CC The invention relates to primers for synthesising full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesising the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesised by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a polypeptide  
CC encoded by a full length human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.  
XX SQ Sequence 399 AA;

Query Match 68.7%; Score 2123; DB 22; Length 399;  
Best Local Similarity 99.7%; Pred. No. 1.5e-204;  
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 178 MSYSVKKKSDVKPFEDFDKYLDPSFQHRHWFSEIFNSFMVIFLVGLYSMLMRTLK 237  
Db 1 MSYSVKKKSDVKPFEDFDKYLDPSFQHRHWFSEIFNSFMVIFLVGLYSMLMRTLK 60

QY 238 DYARYSKEEEMDDMDRLGDGEYQKQVHGVDVFRPSSHPLIFSSSLIGSGCQIFAVSLIVII 297  
Db 61 DYARYSKEEEMDDMDRLGDGEYQKQVHGVDVFRPSSHPLIFSSSLIGSGCQIFAVSLIVII 120

QY 298 VAMEDLYTERGSMSTAIFVYAATSPVNGYFGSLYAROGGRWIKOMFIFGLIPAMV 357  
Db 121 VAMEDLYTERGSMSTAIFVYAATSPVNGYFGSLYAROGGRWIKOMFIFGLIPAMV 180

QY 358 CGTAFFINFIAIYYHASRAIPFGTMAVCCICFVFILPLNLVGTILGRNLSGQNPFCRV 417  
Db 181 CGTAFFINFIAIYYHASRAIPFGTMAVCCICFVFILPLNLVGTILGRNLSGQNPFCRV 240

QY 418 NAVPRPIPEKKWMEPAVIVCLGILPFGSIFEMFYFTSFWAYKTYIYVYGFMMVLVI 477  
Db 241 NAVPRPIPEKKWMEPAVIVCLGILPFGSIFEMFYFTSFWAYKTYIYVYGFMMVLVI 300

QY 478 LCIVTCVTIVCTYFLNAEDYRWQWTSFLSAATAIYVYMYSFYFFKTKMYGLPQTS 537  
Db 301 LCIVTCVTIVCTYFLNAEDYRWQWTSFLSAATAIYVYMYSFYFFKTKMYGLPQTS 360

QY 538 FYFGYMAVFTALGIMCGAIGYMGTSFAVRKIYTNVKID 576  
Db 361 FYFGYMAVFTALGIMCGAIGYMGTSFAVRKIYTNVKID 399

RESULT 9  
AAB75546  
ID AAB75546 standard; Protein; 360 AA.  
XX AAB75546;  
XX AAB75546;  
DT 06-APR-2001 (first entry)  
DE Human secreted protein sequence encoded by gene 41 SEQ ID NO:100.  
KW Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;  
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;  
KW neotropic; neuroprotective; antibacterial; virulicide; fungicide;  
KW ophthalmological; vulnary; autoimmune disease; cardiovascular disorder;  
KW hyperproliferative disorders; cerebrovascular disorder; wound healing;  
KW nervous system disorder; ocular disorder; skin ageing; chemotaxis;  
KW food additive.  
OS Homo sapiens.  
XX  
PN WO200077026-A1.  
XX  
PD 21-DEC-2000.  
XX

PF 01-JUN-2000; 2000MO-US14973.  
XX  
PR 11-JUN-1999; 99US-0138630.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX (ROSE/) ROSEN C A.  
PI Rosen CA, Ruben SM, Komatsoulis GA;  
XX  
DR WPI; 2001-071258/08.  
DR N-PSDB; AAF64216.  
XX  
PT Nucleic acid molecules encoding human secreted proteins, used in  
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and  
PT Parkinson's diseases and cancers  
XX  
XX Claim 11; Page 487-488; 542pp; English.  
XX  
CC Human secreted proteins AAB75506 - AAB75554 are encoded by polynucleotide  
CC sequences AAF64176 - AAF64224. The specification includes amino acid  
CC sequences AAB75555 - AAB75606 which represent fragments of the human  
CC secreted proteins, and protein sequences with which they share homology.  
CC The proteins and polynucleotides, their agonists and antagonists have  
CC activities dependent on the tissues and cells in which they are  
CC expressed, examples of these activities include, immunosuppressive;  
CC antiarthritic; antirheumatic; antiproliferative; cytostatic; cardiant;  
CC vasotropic; cerebroprotective; neotropic; neuroprotective; antibacterial;  
CC virulicide; fungicide; ophthalmological; and vulnary. The proteins,  
CC polynucleotides, agonists and antagonists can be used to treat or detect  
CC or diagnose various diseases and disorders including, autoimmune  
CC diseases e.g. rheumatoid arthritis, hyperproliferative disorders  
CC e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infectious caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin ageing due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities. Included in the invention are  
CC polynucleotide sequences AAF64167 - AAF64175 and peptide AAB75505 which  
CC are used in the isolation, identification and characterisation of the  
CC proteins of the invention.  
XX  
SQ Sequence 360 AA;

Query Match 61.6%; Score 1902; DB 22; Length 360;  
Best Local Similarity 100.0%; Pred. No. 2.1e-182;  
Matches 359; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 218 MMVIFLVGLYSMLMRTLKDYARYSKEEEMDDMDRLGDGEYQKQVHGVDVFRPSSHPLI 277  
Db 1 MMVIFLVGLYSMLMRTLKDYARYSKEEEMDDMDRLGDGEYQKQVHGVDVFRPSSHPLI 60

QY 278 FSSSLIGSGCQIFAVSLIIVIAMIEDLYTERGSMSTAIFVYAATSPVNGYFGSLYARQ 337  
Db 61 FSSSLIGSGCQIFAVSLIIVIAMIEDLYTERGSMSTAIFVYAATSPVNGYFGSLYARQ 120

QY 338 GGRRWIKOMFIFGLIPAMVCGTAFFINFIAIYYHASRAIPFGTMAVCCICFVFILPLN 397  
Db 121 GGRRWIKOMFIFGLIPAMVCGTAFFINFIAIYYHASRAIPFGTMAVCCICFVFILPLN 180

QY 398 LVGTILGRNLSGQNPFCRVNAVPRPIPEKKWMEPAVIVCLGILPFGSIFEMFYFT 457  
Db 181 LVGTILGRNLSGQNPFCRVNAVPRPIPEKKWMEPAVIVCLGILPFGSIFEMFYFT 240

QY 458 SFWAYKIYYVYGFMMVLVLICIVTCVTIVCTYFLNAEDYRWQWTSFLSAATAIYVY 517  
Db 241 SFWAYKIYYVYGFMMVLVLICIVTCVTIVCTYFLNAEDYRWQWTSFLSAATAIYVY 300

QY 518 MYSFYFFKTKMYGLFQTSFYFGYMAVFTALGIMCGAIGYMGTSFAVRKIYTNVKID 576  
|||||

Db 301 MYSEYFFFTKMYGLFQTSFYFGYNAVFSTALGIMCGAIGYMGTSAPVRKIYTNWKID 359

RESULT 10  
AAM93524

ID AAM93524 standard; Protein; 329 AA.

XX AC AAM93524;

XX AC

XX XX 06-NOV-2001 (first entry)

XX XX Human polypeptide, SEQ ID NO: 3257.

XX XX Human; full length cDNA; cDNA synthesis; oligo-capping.

XX KW Homo sapiens.

XX OS

XX PN EP1130094-A2.

XX XX

XX PD 05-SEP-2001.

XX PF

XX PF 07-JUL-2000; 2000EP-0114089.

XX PF

XX PR 08-JUL-1999; 99JP-0194486.

XX PR 11-JAN-2000; 2000JP-0118774.

XX PR 02-MAY-2000; 2000JP-0183765.

XX XX

PA (HELI-) HELIX RES INST.

XX XX

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;  
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;  
XX XX

DR WPI: 2001-524255/58.  
DR N-PSDB; AAK94454.

XX XX

XX 830 Primers useful for synthesizing full length cDNA clones and their  
PT use in genetic manipulation -  
PT XX

PS Claim 8; SEQ ID NO 3257; 1380pp + sequence listing; English.

XX XX

CC The invention relates to primers for synthesizing full length cDNA  
CC clones. 830 cDNA molecules encoding a human protein have been  
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA  
CC molecules have been determined. Primers for synthesizing the full length  
CC cDNA are useful for clarifying the function of the protein encoded by  
CC the cDNA. The full length clones were obtained by construction of full  
CC length enriched cDNA libraries that were synthesized by the oligo-capping  
CC method. The primers enable the production of the full length cDNA easily  
CC without any special methods. The present sequence is a polypeptide  
CC encoded by a full length human cDNA of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in CD-ROM format directly from EPO.

XX XX

SQ Sequence 329 AA:

Query Match 56.9%; Score 1757; DB 22; Length 329;  
Best Local Similarity 100.08; Pred. No. 7.1e-168;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 248 MDDMDRLDGEYGWKQVHGDVFRPSSHPLFSSLLIGSGCIFAVSLIIVIAMIEDLYTE 307  
|||||  
Db 1 MDDMDRLDGEYGWKQVHGDVFRPSSHPLFSSLLIGSGCIFAVSLIIVIAMIEDLYTE 60

QY 308 RGSMLSTAIFVVAATSPVNGYFGGSLYARQGGRRWIKOMFTGAFILPAMVCGTAFFINFI 367  
|||||  
Db 61 RGSMLSTAIFVVAATSPVNGYFGGSLYARQGGRRWIKOMFTGAFILPAMVCGTAFFINFI 120

QY 368 AIYXHASRAIPGTMWAVCCICFFVILPLNLVGTILGNLSGQNFPCRNVAPRPIPEK 427  
|||||

Db 121 AIYXHASRAIPGTMWAVCCICFFVILPLNLVGTILGNLSGQNFPCRNVAPRPIPEK 180

QY 428 KWFMEPAIVCGGILPFGSGIEMFYFTSPWAKYIVYVGFMMVLVLITLCITVTCVTI 487  
|||||

[illegible]



KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 99US-0121825.  
XX 05-MAR-1999; 99US-0123180.  
XX 09-MAR-1999; 99US-0123548.  
XX 23-MAR-1999; 99US-0125788.  
XX 25-MAR-1999; 99US-0126264.  
XX 29-MAR-1999; 99US-0126785.  
XX 01-APR-1999; 99US-0127462.  
XX 06-APR-1999; 99US-0128234.  
XX 08-APR-1999; 99US-0128714.  
XX 16-APR-1999; 99US-0129845.  
XX 19-APR-1999; 99US-0130077.  
XX 21-APR-1999; 99US-0130449.  
XX 23-APR-1999; 99US-0130510.  
XX 28-APR-1999; 99US-0130891.  
XX 30-APR-1999; 99US-0131449.  
XX 30-APR-1999; 99US-0132048.  
XX 04-MAY-1999; 99US-0132407.  
XX 05-MAY-1999; 99US-0132484.  
XX 06-MAY-1999; 99US-0132485.  
XX 06-MAY-1999; 99US-0132486.  
XX 07-MAY-1999; 99US-0132863.  
XX 11-MAY-1999; 99US-0134256.  
XX 14-MAY-1999; 99US-0134218.  
XX 14-MAY-1999; 99US-0134219.  
XX 14-MAY-1999; 99US-0134221.  
XX 18-MAY-1999; 99US-0134370.  
XX 19-MAY-1999; 99US-0134768.  
XX 20-MAY-1999; 99US-0134941.  
XX 21-MAY-1999; 99US-0135124.  
XX 24-MAY-1999; 99US-0135353.  
XX 25-MAY-1999; 99US-0135629.  
XX 27-MAY-1999; 99US-0136021.  
XX 28-MAY-1999; 99US-0136392.  
XX 01-JUN-1999; 99US-0136782.  
XX 03-JUN-1999; 99US-0137222.  
XX 04-JUN-1999; 99US-0137528.  
XX 07-JUN-1999; 99US-0137502.  
XX 08-JUN-1999; 99US-0137724.  
XX 10-JUN-1999; 99US-0138094.  
XX 10-JUN-1999; 99US-0138540.  
XX 14-JUN-1999; 99US-0138847.  
XX 16-JUN-1999; 99US-0139119.  
XX 16-JUN-1999; 99US-0139452.  
XX 17-JUN-1999; 99US-0139453.  
XX 18-JUN-1999; 99US-0139492.  
XX 18-JUN-1999; 99US-0139454.  
XX 18-JUN-1999; 99US-0139455.  
XX 18-JUN-1999; 99US-0139456.  
XX 18-JUN-1999; 99US-0139457.  
XX 18-JUN-1999; 99US-0139458.  
XX 18-JUN-1999; 99US-0139459.  
XX 18-JUN-1999; 99US-0139460.  
XX 18-JUN-1999; 99US-0139461.  
XX 18-JUN-1999; 99US-0139462.  
XX 18-JUN-1999; 99US-0139463.  
XX 18-JUN-1999; 99US-0139750.  
XX 18-JUN-1999; 99US-0139763.  
XX 21-JUN-1999; 99US-0139817.  
XX 22-JUN-1999; 99US-0139899.  
XX 23-JUN-1999; 99US-0140353.  
XX 23-JUN-1999; 99US-0140354.  
XX 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.





PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144009.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 20-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145089.  
PR 23-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147204.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147192.  
PR 05-AUG-1999; 99US-0147260.  
PR 06-AUG-1999; 99US-0147303.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 12-AUG-1999; 99US-0148341.  
PR 13-AUG-1999; 99US-0148565.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149368.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149723.  
PR 20-AUG-1999; 99US-0149929.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151080.  
PR 30-AUG-1999; 99US-0151303.  
PR 31-AUG-1999; 99US-0151438.  
PR 01-SEP-1999; 99US-0151930.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 15-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154039.  
PR 20-SEP-1999; 99US-0154779.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155659.  
PR 28-SEP-1999; 99US-0156458.  
PR 29-SEP-1999; 99US-0156596.  
PR 04-OCT-1999; 99US-0157117.  
PR 05-OCT-1999; 99US-0157753.  
PR 06-OCT-1999; 99US-0157865.  
PR 07-OCT-1999; 99US-0158029.  
PR 08-OCT-1999; 99US-0158232.  
PR 12-OCT-1999; 99US-0158369.  
PR 13-OCT-1999; 99US-0159293.  
PR 13-OCT-1999; 99US-0159294.  
PR 13-OCT-1999; 99US-0159295.  
PR 14-OCT-1999; 99US-0159329.  
PR 14-OCT-1999; 99US-0159330.  
PR 14-OCT-1999; 99US-0159331.  
PR 14-OCT-1999; 99US-0159637.  
PR 14-OCT-1999; 99US-0159638.  
PR 18-OCT-1999; 99US-0159584.  
PR 21-OCT-1999; 99US-0160741.  
PR 21-OCT-1999; 99US-0160767.  
PR 21-OCT-1999; 99US-0160768.  
PR 21-OCT-1999; 99US-0160770.  
PR 21-OCT-1999; 99US-0160814.  
PR 21-OCT-1999; 99US-0160815.  
PR 22-OCT-1999; 99US-0160980.  
PR 22-OCT-1999; 99US-0160981.  
PR 22-OCT-1999; 99US-0160989.  
PR 25-OCT-1999; 99US-0161404.  
PR 25-OCT-1999; 99US-0161405.  
PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161992.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.

Query Match 54.0%; Score 1669; DB 21; Length 583;  
Best Local Similarity 54.4%; Pred. No. 1.1e-158;  
Matches 314; Conservative 106; Mismatches 145; Indels 12; Gaps 5;

QY 5 LLLLLPTRADEHHTVODKEEVVLMNNTVGPYHNROETKYKYSPLPCVSGSKKSISHYH 64  
DB 14 LLVSLTPTTFASDSDBKQABEQVTLWNKVGYPNPOETNYSLPFCRPSGNNV-HKW 72  
QY 65 ETLEALOGVELESGLDIKEDDVPATYCEIDLDEKRDFAVYAIKNHXYQWYIDDL 124  
DB 73 GGLGVGGLNELIDSEIAIKMKNVRSVIPCLEDEAKVHKHFDKDAESSWFEFFM--- 129  
QY 125 PIWGTIVGE--ADENGED--YVLWTYKKEIGFNGNRIVDVNLTSEKGVKLVNPKIQMSY 180  
DB 130 ---GFVGEHPDKNSCKHVLVTHKNVVKYKNDQIIHVNLTQDNPRPLEAGKKMDLTY 186  
QY 181 SVKWKSDVKFEDRPDKYLDSPFOHRTHWFSIFNSFMVIFLVGLVSMILMRTLRLKDYA 240  
DB 187 SVQWITPTNVTFARRPDVLDYDPFEHQTHWFSIFNSFMVIFLVGLVSMILMRTLRLDYA 246  
QY 241 RYSKE-BEMDDMDRLDGEYQKQVHGDFRPSHPLIFSSILGSGCOIFAVSLIIVIA 299  
DB 247 KYAREDDLESRLDVSSESGKLVHGDFRPASSVLVLSAVGTGAOLALLVLLVILMA 306  
QY 300 MIEDLYTERGSMLETAIFVYAATSPVNGYFGGSLYAROGGRWIKQMFIGAFLIPAMVCG 359  
DB 307 IGVTLVGRGALVITTFIVCYALTSPVSGYVSGMGYSRSGGKHWIKCMVLTASLPFLCFG 366  
QY 360 TAFFINTAIYHASRAIPFGTMVAVCCICFVILPLNLVTILGRNLGSGQPNFPCRVNA 419  
DB 367 IGFLNTTAIFYGLAAIPFGTMVAVFVINGFISFPLALLGTVVGRNWSGAPNRCRVKT 426  
QY 420 VPRPIPEKKWMEPAVIVCLGILPFGSIFTEMVFIETSEWAYKIYVYVGFMLVLVILC 479  
DB 427 IPRPIPEKKWLTSPVSLMGLLPFGSIFTEMVFIETSEWAYKIYVYVGFMLLVFVILV 486  
QY 480 IVTCVTIVCTYFLLNAEDYRWQWTSFSLAASATAIYVYVYFYFFKTKMYGLFQTSFY 539  
DB 487 IVTCVTIVCTYFLLNAENYHWQWTSFSLAASATAIYVYVYFYFFKTKMYGLFQTSFY 546  
QY 540 FGYMAVFTALGIMCGAIGYWGTSFAVRKRIYTNVKID 576  
DB 547 FGYTMFGLGLGILCGAVGLGSLNLFVRRIRYRNKCD 583

RESULT 15  
AAG45637  
XX AAG45637 standard; Protein: 602 AA.  
XX AC AAG45637;  
XX DT 18-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 57320.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-0301439.  
XX PR 25-FEB-1999; 99US-0121825.  
XX PR 05-MAR-1999; 99US-0123180.  
XX PR 09-MAR-1999; 99US-0123548.  
XX PR 23-MAR-1999; 99US-0125788.  
XX PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 23-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.



Search completed: January 18, 2003, 20:53:33  
Job time : 43 secs

2003-01-23